



Human Tooth Crack Image Analysis with Multiple Deep Learning Approaches

Zheng Li¹ · Zhongqiang Li¹ · Ya Zhang¹ · Huaizhi Wang¹ · Xin Li² · Jian Zhang³ · Waleed Zaid⁴ · Shaomian Yao⁵ · Jian Xu¹

Received: 17 May 2024 / Accepted: 2 September 2024
© The Author(s) under exclusive licence to Biomedical Engineering Society 2024

Abstract

Tooth cracks, one of the most common dental diseases, can result in the tooth falling apart without prompt treatment; dentists also have difficulty locating cracks, even with X-ray imaging. Indocyanine green (ICG) assisted near-infrared fluorescence (NIRF) dental imaging technique can solve this problem due to the deep penetration of NIR light and the excellent fluorescence characteristics of ICG. This study extracted 593 human cracked tooth images and 601 non-cracked tooth images from NIR imaging videos. Multiple imaging analysis methods such as classification, object detection, and super-resolution were applied to the dataset for cracked image analysis. Our results showed that machine learning methods could help analyze tooth crack efficiently: the tooth images with cracks and without cracks could be well classified with the pre-trained residual network and squeezenet1_1 models, with a classification accuracy of 88.2% and 94.25%, respectively; the single shot multi-box detector (SSD) was able to recognize cracks, even if the input image was at a different size from the original cracked image; the super-resolution (SR) model, SR-generative adversarial network demonstrated enhanced resolution of crack images using high-resolution concrete crack images as the training dataset. Overall, deep learning model-assisted human crack analysis improves crack identification; the combination of our NIR dental imaging system and deep learning models has the potential to assist dentists in crack diagnosis.

Keywords NIRF dental imaging · Human tooth cracks diagnosis · Crack detection · Deep learning

Introduction

Dental health has drawn significant attention nowadays since dental disorders, such as caries (decay), fractures, and impacted wisdom teeth make a massive number population suffer pain and uncomfortable. Caries lesions and decays affect most of the population; wisdom teeth influence up to half population [1]. Craze line, fractured cusp, vertical root fracture, cracked tooth, and split tooth are the five common types of tooth fractures [2]; the craze line is a fracture that happens only outside the crown (enamel), which is currently deemed not harmful to dental health [3]. However, these small cracks may lead to pulp necrosis if they grow larger [4]. In addition, 34% – 74% of adults suffer sound tooth structure loss caused by their cracked teeth [5]. Periodic examination and early detection of cracked teeth are therefore essential to decrease the possibility of crack growth and consequential deterioration of the teeth. The most common dental imaging method is X-ray imaging, such as 2D radiograph and 3D cone beam computed tomography (CBCT)

Associate Editor Eiji Tanaka oversaw the review of this article.

✉ Jian Xu
jianxu1@lsu.edu

- ¹ Division of Electrical and Computer Engineering, College of Engineering, Louisiana State University, Baton Rouge, LA 70803, USA
- ² Section of Visual Computing and Creative Technology, School of Performance, Visualization, & Fine Art, Texas A & M University, College Station, TX 77843, USA
- ³ Division of Computer Science & Engineering, College of Engineering, Louisiana State University, Baton Rouge, LA 70803, USA
- ⁴ Oral and Maxillofacial Surgery, School of Dentistry, Louisiana State University Health Science Center, Baton Rouge, LA 70808, USA
- ⁵ Department of Comparative Biomedical Science, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA 70803, USA

[6]. However, X-ray dental imaging has its disadvantages: (1) it generates ionizing radiation, which may be harmful to human health (e.g., low birth weight baby, and induction of cancer) [7]; and (2) the crack on the enamel cannot be identified by micro-computerized tomography (micro-CT) [1]. Other modern imaging techniques, such as optical coherence tomography (OCT) and ultrasound, are also employed in dental imaging research but are less efficient [8, 9]. Though the swept-source OCT can demonstrate the tomography between tooth enamel and dentin, several disadvantages limit its application: (1) the patients are required to be in the same position to achieve good imaging quality for a full mouth lesion detection, which is insufficient; (2) the shallow penetration of current swept source OCT is ~ 3 mm, which needs to be improved for deep crack detection [10]. Few studies of ultrasound have been investigated for tooth crack detection because of the low resolution; the new ultrasonic technology, laser ultrasound, can detect tooth crack while the laser radiation may damage oral tissue [11, 12]. Magnetic resonance imaging (MRI) can identify tooth caries and cracks but is normally expensive and requires massive scanning time [13].

Infrared imaging, especially near-infrared (NIR) attention in the tooth imaging field recently [14]. Indocyanine green (ICG), the U.S. Food and Drug Administration (FDA) approved medical imaging dye, assisted-NIR imaging is an efficient and safe method in dental disease diagnosis; this system is also capable of capturing cracks on the enamel [15, 16]. NIR imaging can be separated into two categories regarding wavelength: NIR-I (wavelength between 700 nm to 950 nm) and NIR-II (wavelength between 1000 nm to 1700 nm) [17, 18]. Although our NIRF imaging scheme has demonstrated superior imaging of cracks, including those often missed by the current dental X-ray/CT, it is time-consuming for dentists to manually read each frame of video to find all important cracks [14, 17, 19, 20]. Nowadays, the main method of crack detection in dentistry is visible light reflection/transmission with dentists' naked eyes. Some researchers stated that dental X-ray/CT, particularly CBCT, is the gold standard method for dental imaging [21]. However, cracks can only be detected when it is larger than 100 μm , as shown in our previous studies and other literature [1, 22, 23]. In addition, the micro-CT is not clinically reasonable because of the high radiation. The NIR dental image and video enhancement and analysis are needed to improve the efficiency of the process and make it practical for dentists.

Machine learning-aided biomedical image processing has been fast developing since the improvement of computation in recent years [24]. Kist and Döllinger implemented an edge tensor processing units (TPUs) based biomedical imaging segmentation approach, which shortened the waiting time for patients [25]. Brattain et al. summarized the status,

approaches, and future of machine learning-assisted medical ultrasound (US) techniques [26]. Bloice et al. reported the Augmentor package (designed to assist the image augmentation and generation for machine learning tasks) of python for biomedical image augmentation purposes [27]. Deep learning methods could help in reconstructions and enhancements of the optical microscopy images for optimizing disease diagnosis, such as reducing imaging system cost (e.g., super-resolution) and improving accuracy (e.g., virtual staining) [28, 29]. On the other hand, biomedical signal classification and prediction play a significant role in diagnosing diseases [30]. In addition, object detection can locate damaged structures and cell nuclei after training [31, 32]. Single-shot multi-box detection (SSD) is a popular and accurate object detection model. Rashid et al. applied the SSD and level set segmentation method to segment melanoma lesions [33]. The super-resolution (SR) technique becomes necessary since the resolution of many medical images is quite low [34]. Chen and colleagues applied deep connected neural networks (NN) for brain magnetic resonance imaging (MRI) SR images, their results outperformed other existing deep learning algorithms due to the densely connected NN [35].

Most of the current deep learning-aided dental analyses use radiographs, such as those produced by computed tomography (CT). Lee et al. applied a U-shaped based convolutional neural network (CNN) for caries detection on the bitewing radiograph dataset, which improved the dentist's diagnostic performance [36]. Choi's group collected 3000 premolar and molar periapical radiographs and implemented a binary CNN model to classify caries images and non-caries images, achieving an average accuracy of $\sim 85\%$ from three different molar models [37]. Jaskari and colleagues developed a machine-learning model for automatically locating the mandibular canal based on the CBCT volume data [38]. To the best of our knowledge, no deep-learning-based NIR-assisted dental imaging technique for cracked and non-cracked tooth image analysis has been reported to this date. In our previous studies, we demonstrated that ICG-assisted NIRF dental imaging was capable of identifying dental diseases, such as caries and cracks; the system had the potential to be an efficient and safe tool for dental examinations [22, 39, 40]. In this work, we first built a human tooth image dataset that included tooth images with and without cracks from NIR tooth videos. This dataset is also innovatively analyzed with multiple machine learning methods for accurate diagnosis of tooth cracks, including cracked and non-cracked images classification, crack identification, and high-resolution crack images.

Materials and Methods

Crack Tooth Image Collection with NIRF Dental Imaging System

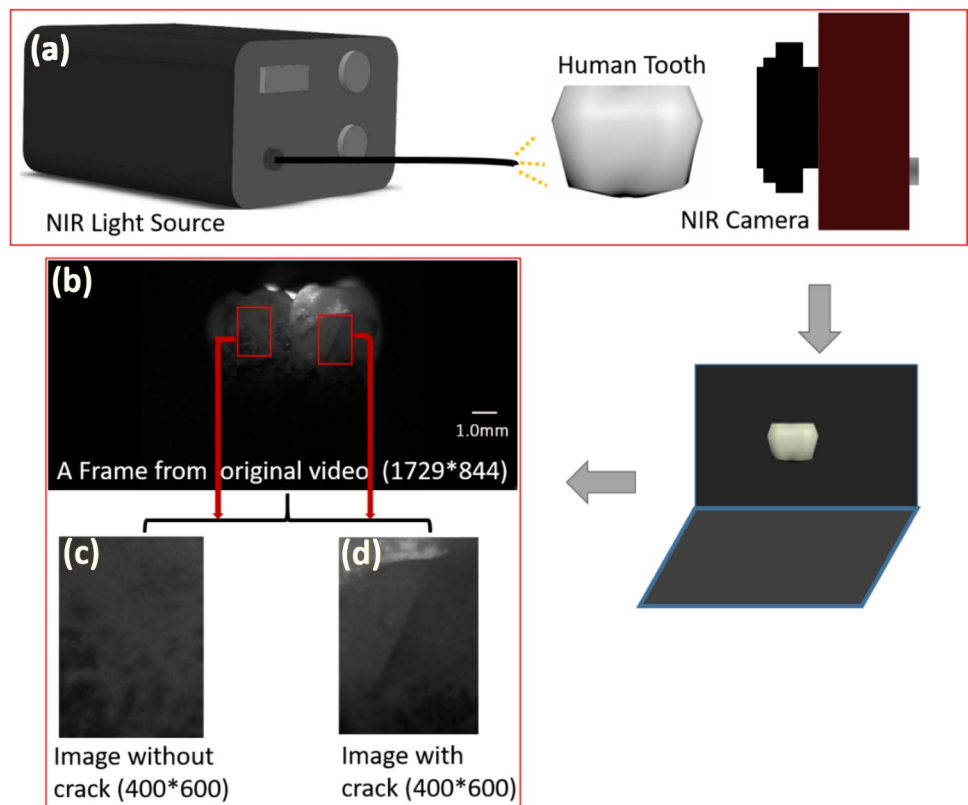
The human teeth were from Louisiana State University Health Science Center (LSUHSC) (Baton Rouge, USA); a dentist with more than 10 years of experience and his residents helped collect the teeth and confirm the cracks using visible light with their naked eyes. Tooth videos were taken by the NIR imaging system (Fig. 1a); the imaging setup could be found in our previous studies [19, 22]. Briefly, the light from the 785 nm laser (Turkey Raman Lasers 785, Ocean Optics) went through a 785 nm band pass filter (Thorlabs) by an optical fiber; the extracted human teeth were immersed in ICG solution (concentration: 50 μ M; duration of immersion: 1 min, 10 min, 4 h, 24 h), then shined with NIR light (both transillumination and reflection), and recorded by NIR cameras (NIR-I camera, Mako U130B, Allied Vision; NIR-II camera, Goldeye G008, Allied Vision); the cracks on the tooth could be found and collected in the videos. Figure 1b showed an original frame of the video (image size: 1729×844 , all the units are pixels); the images without a crack (Fig. 1c) and with a crack (Fig. 1d) were trimmed to the size of 400×600 from the original image by a lab-developed

software (MATLAB, R2019a; MathWorks Inc, Natick, Mass). In this tool, we could select the crack in any size and generate the 400×600 image by interpolation. The dataset contained 593 crack tooth images and 601 non-crack tooth images. The Institutional Review Board of Louisiana State University approved the experimental procedures (IRB#E11061).

Classification of Tooth Images With Crack and Without Crack

Two deep neural networks were implemented and compared for cracked image and non-cracked image classification. The first network was the ImageNet pre-trained residual neural network (ResNet50) model (Fig. 2a). This model included five stages: stage1 consisted of a convolutional layer, a batch normalization layer, a ReLU layer, and a max-pooling layer; stage2 and stage5 had one convolutional block (each block contained three convolutional layers) and two identity blocks (each block also contained three convolutional layers); stage3 had one convolutional block and three identity blocks; stage4 had one convolutional block and five identity blocks; all the weights were frozen. Two fully connected layers were added: the first one had 128 neurons, and the second layer had two neurons for the classification [41, 42]. 80% of the tooth image data (~474 cracked tooth images and 480 non-cracked tooth images) were used for training and 20% (~119 cracked tooth

Fig. 1 Tooth image with crack and without crack collection from NIRF dental imaging system



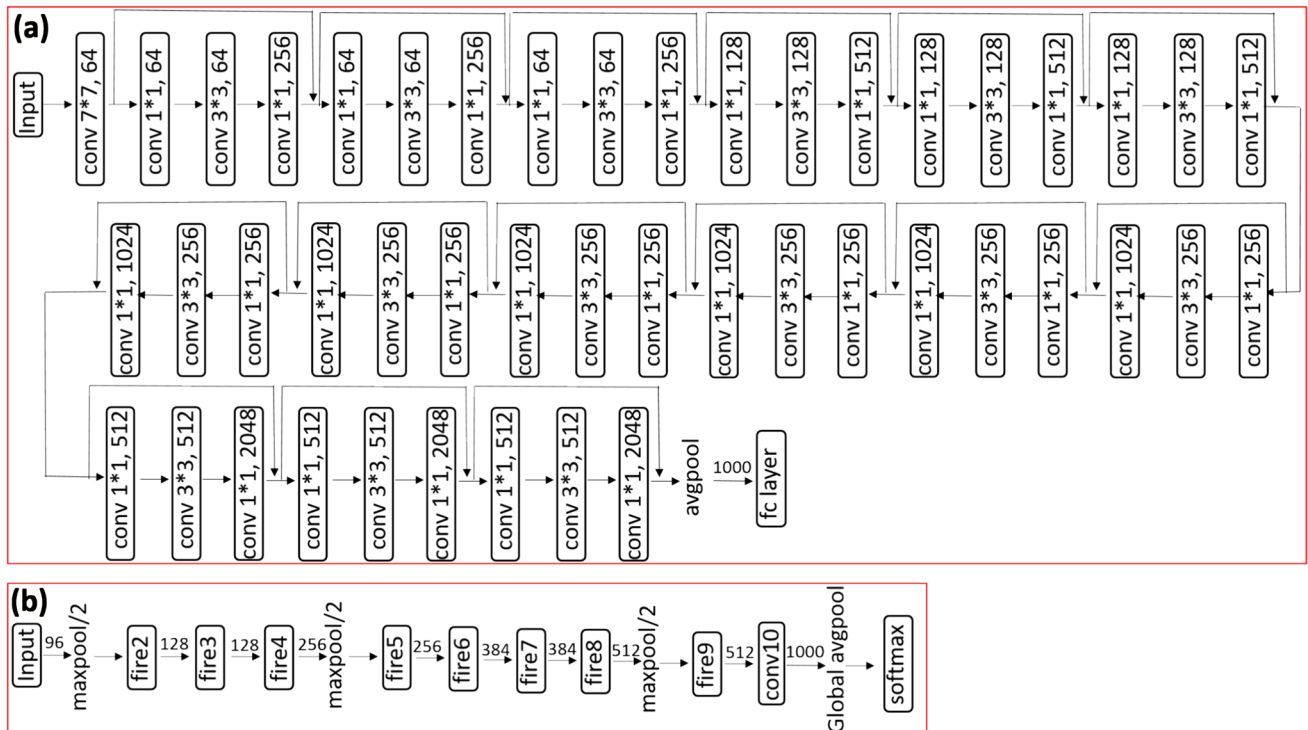


Fig. 2 The network architecture of classification models. **a** ResNet50; **b** squeezenet1_1

images and 121 non-cracked tooth images) for testing. The other network was squeezenet1_1 (Fig. 2b), which was also a pre-trained model based on ImageNet; this net consisted of two convolutional layers, 8 fire layers, and a final classifier layer [43]. ~83% of the data (1000 tooth images) were used for training and validation, and ~17% (193 tooth images) for testing.

Crack Detection of the Tooth Crack Image

Object detection was an excellent tool to locate the target in the image or video. In this study, we used the SSD model to detect and locate the tooth crack from the crack images. The SSD model was based on VGG16, which had 16 parameterized layers (13 convolutional layers and 3 fully connected layers) [44]. We created and labeled 378 tooth crack images as a new class since there was no crack label in the existing dataset; the labeled crack images were trained together with the Microsoft common objects in context (COCO) dataset and the PASCAL visual object classes challenge (VOC2007) dataset [45, 46]. The cracks were labeled by the labelImg [47].

Tooth Crack Super-Resolution with Concrete Crack Images

Super-resolution ResNet (SRResNet) and super-resolution generative adversarial network (SRGAN) were

implemented for improving the resolution of the tooth crack images. The SRResNet contained one convolutional module, 16 residual modules, one convolutional module, two sub-pixel convolutional modules, and one convolutional module, sequentially [48]. The SRGAN included a Generator and a Discriminator: the Generator was the same network structure as SRResNet; the Discriminator was a classification model to identify whether the image was the generated image or the original image, which contained a convolutional layer, Leaky ReLU layer, seven convolutional modules, dense layer, Leaky ReLU layer, and Sigmoid layer [48]. In this work, we used high-resolution concrete crack images as the training dataset to mimic high-resolution tooth crack images. In addition, the concrete images were large and distinct. The concrete crack dataset had 20,000 cracked images and 20,000 non-cracked images; each image size was 227×227 [49]. The test image was a low-resolution tooth crack image. The SRResNet model was trained for 100 epochs; the SRGAN model was trained for 50 epochs.

The machine learning workstation is based on an HP Z2620 computer (CPU: Intel Xeon processor E5-1600 product family; Memory: 16 GB) with an additional NVIDIA K80 GPU. All the machine-learning approaches were implemented in Python and PyTorch.

Results

Classification Performances of Tooth Images With and Without Crack

The classification accuracy of the pre-trained ResNet50 model was 88.2%. The training and test accuracy became stable at ~6 epochs; the training and test loss was stable at ~5 epochs (Fig. 3a).

The classification performances of squeezenet1_1 outperformed ResNet50, with an accuracy of 94.25%, a sensitivity of 95.0%, and a specificity of 93.49%; the correct predictions of 10 runs were 94.95% for positive images and 93.48% for negative images. The training and test accuracies were stable after 10 epochs; the training loss and test loss became stable after 4 epochs (Fig. 3b).

Crack Identification of the Tooth Crack Image

Three images (one was from the training dataset, and the other two were new images) were tested for crack

identification with the SSD model. Figure 4b–d demonstrated the typical results: Fig. 4a is a ground truth image (the blue box was the manually labeled crack) and this image was also tested as the input to the SSD model (Fig. 4b), the slot captured most of the crack compared to the ground truth; Fig. 4c showed a random crop image from the video frame in Fig. 1 (207×245), the crack was captured successfully; the crack in the original video frame image was also identified (Fig. 4d), only with a larger detector box. The *teethcrack* value is the IoU (intersection over union), which is a metric of object detection; all the IoU values were beyond 0.9, indicating cracks could be well identified.

Tooth crack super-resolution analysis

The loss function of SRResNet was the mean square error (MSELoss), which became stable at ~30 epochs (Fig. 5).

The SRGAN had three losses illustrated in Fig. 6: loss (c), loss (a), and loss (d), which were content loss, generator loss, and discriminator loss, respectively. The losses in SRGAN were not stable since the generator and discriminator were against each other.

Fig. 3 Accuracy and loss to epochs curve of the ResNet (a) and squeezenet1_1 (b)

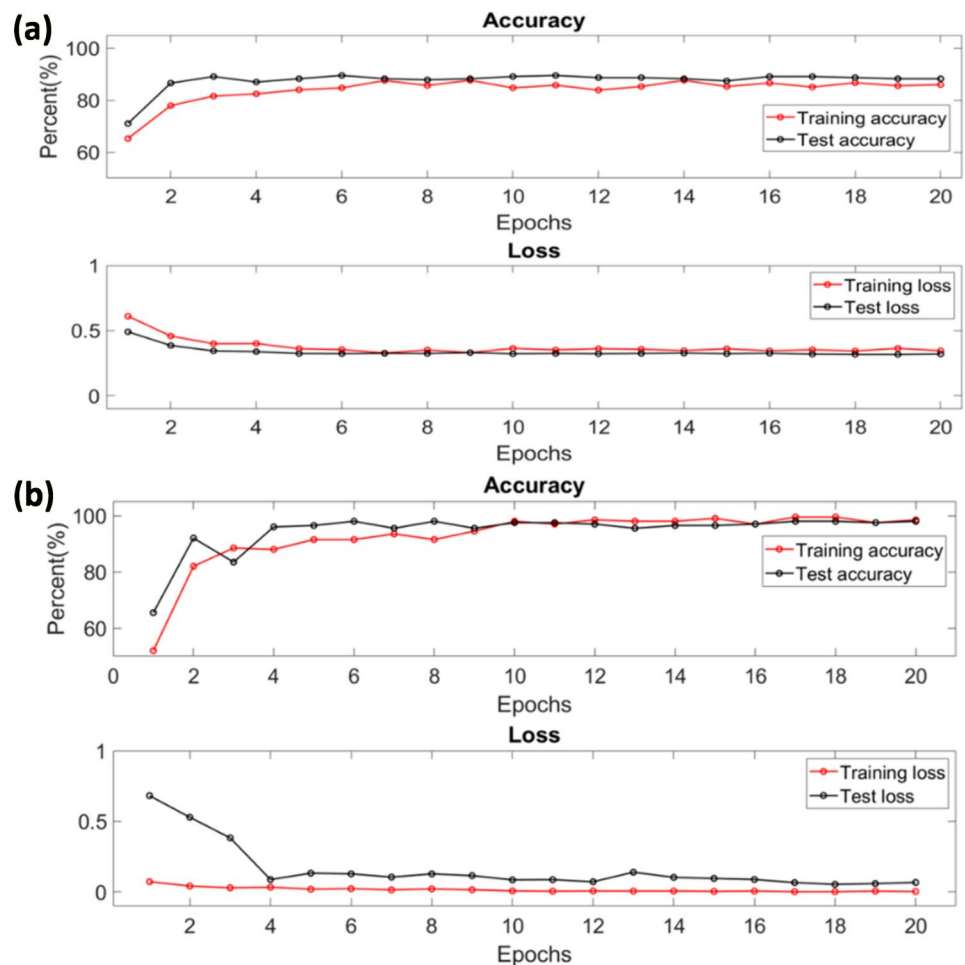


Fig. 4 Tooth crack identification with SSD. **a** ground truth; **b** crack identification of ground truth (**a**); **c** identification of random size cropped image of the video frame in Fig. 1; **d** crack identification of the video frame in Fig. 1. The X–Y axis of b–d demonstrated the image size

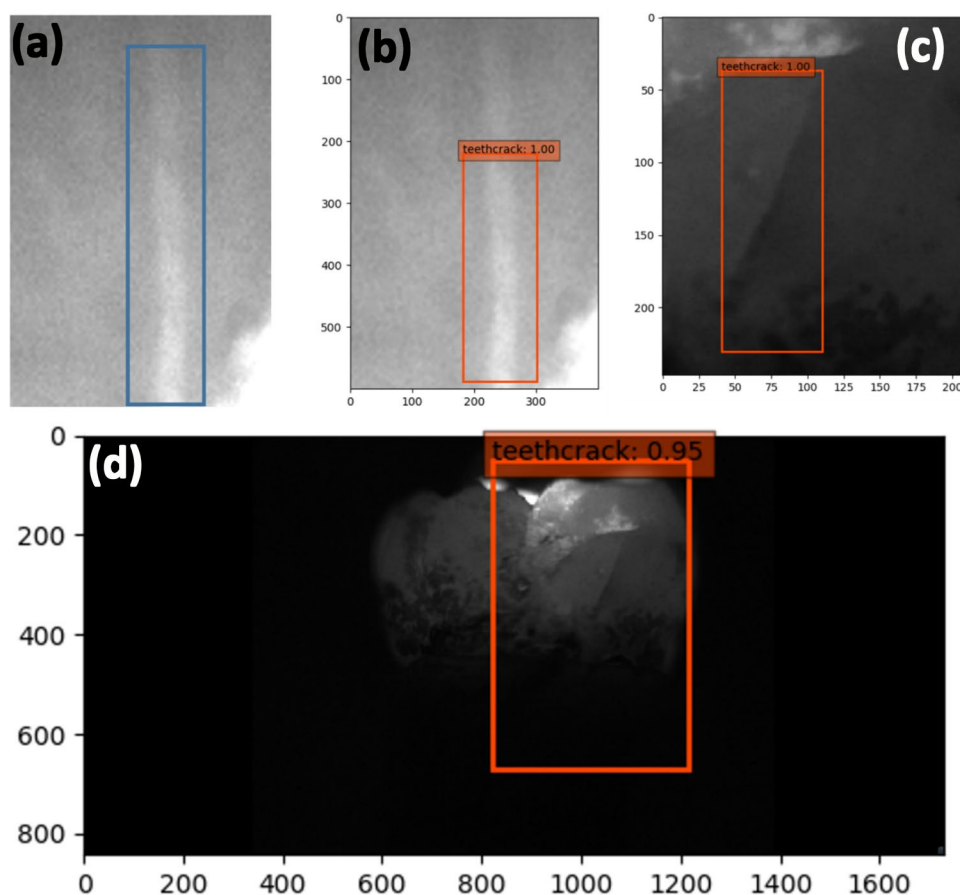


Fig. 5 Loss to epoch curve of SRResNet

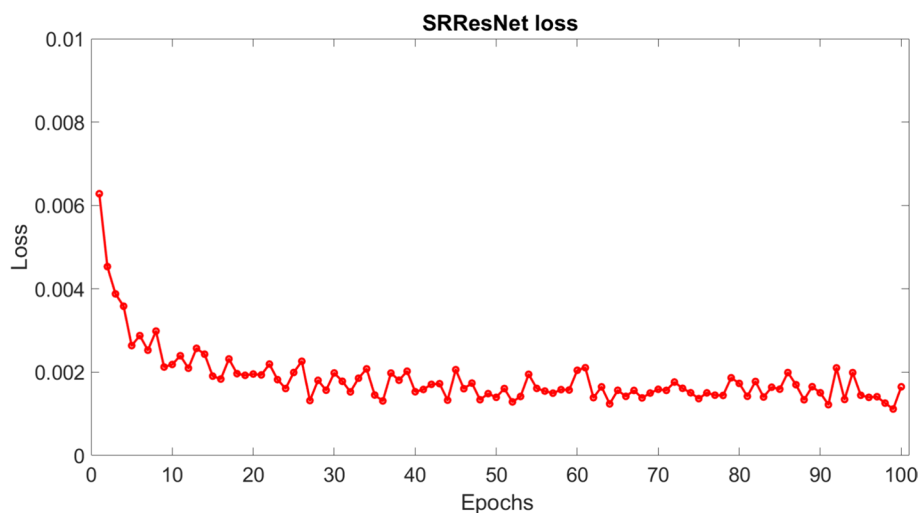


Figure 7 illustrates the concrete crack super-resolution images of SRResNet (epoch 100) and SRGAN (epoch 50) during the training; the first row presents the low-resolution images; the second row presents the super-resolution images; the last row presents the ground truth images. SRGAN images were closer to the ground truth than the SRResNet, which were smoother and more detailed.

We tested our tooth crack images in the SRResNet and SRGAN models to get super-resolution crack images. Figure 8a (400×600 , resolution in pixels) was the original crack image; Fig. 8b (6400×9600) was the image generated

Fig. 6 Loss to epoch curves of SRGAN. Loss c represents contest loss; loss a represents generator loss; loss d represents discriminator loss

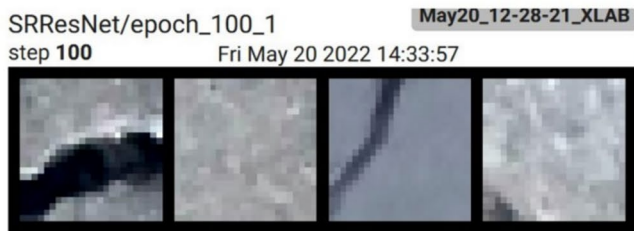
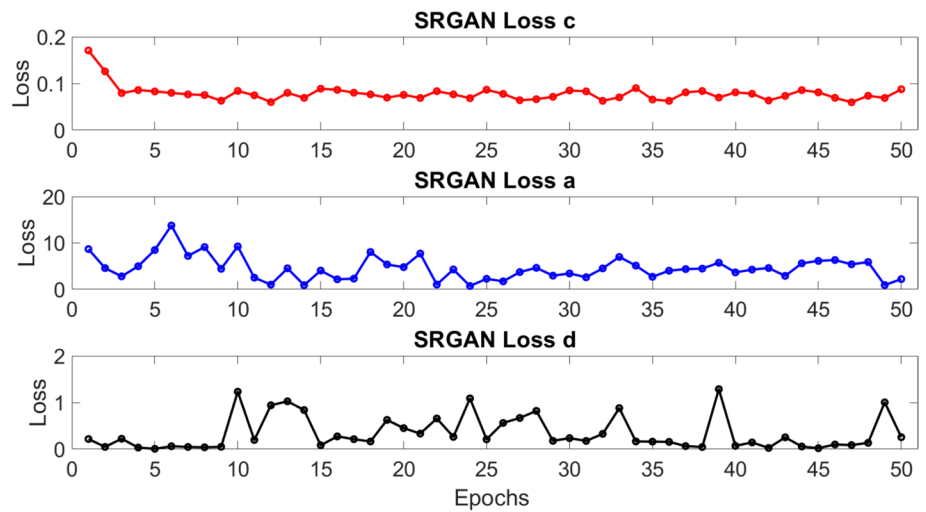


Fig. 7 Concrete super-resolution images with or without cracks in SRResNet and SRGAN models

by the SRResNet model, which was darker than Fig. 8a, but the crack line was identifiable as a clear and black line. Figure 8c (1600 × 2400) was generated by the SRGAN model;

Fig. 8d (6400 × 9600) was generated by the SRGAN model from Fig. 8c, with a higher resolution.

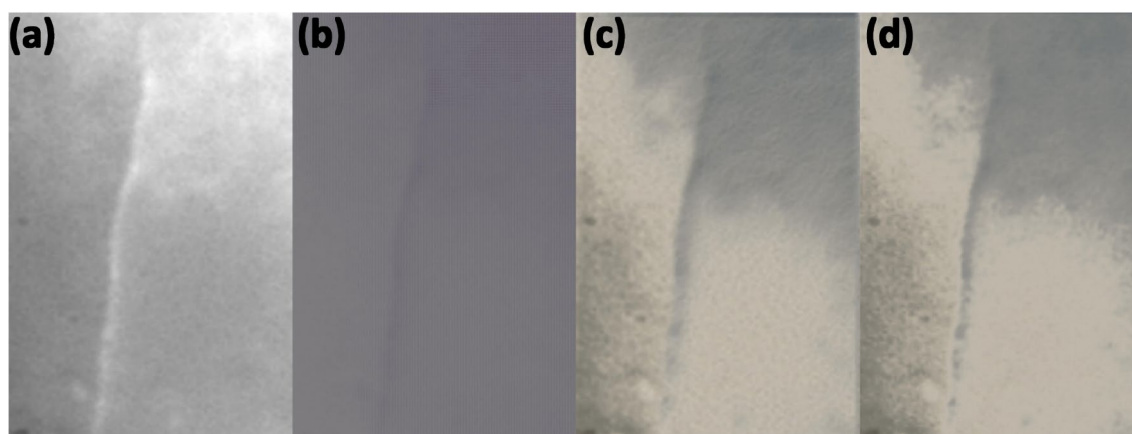


Fig. 8 Tooth crack super-resolution images generated by SRResNet and SRGAN. **a** original image; **b** SRResNet generated crack image; **c** SRGAN generated image; **d** SRGAN generated image from (c)

Discussions

The human tooth enamel is composed of carbonated apatite. Although the enamel is the hardest part of the human body, craze lines may occur because of age, trauma, and grinding. Detecting the small cracks, especially the early-stage craze line, is challenging in dentistry. Our NIR dental imaging system demonstrated that cracks could be detected while X-raying imaging failed. When the NIR light shone on the tooth from different angles, the enamel was bright while the crack was dark according to NIR videos and images. The gap of cracks prevented light from going through, which formed the dark line.

In this work, 593 cracked and 601 non-cracked human tooth images were collected from NIR imaging videos; these images were analyzed with multiple modern machine learning models, such as classification, crack identification, and super-resolution. Our previous studies demonstrated that human tooth cracks on the enamel-dentin can be captured by the NIR dental imaging system, but not by the more commonly utilized micro-CT data. Because some of these cracks were challenging to find and locate, further image and video processing were necessary to optimize our dental imaging system for the dental crack diagnosis.

Binary classification with a pre-trained model was applied to differentiate the images with crack and without crack. One advantage of transfer learning is time efficiency. The parameter and model can be used directly with minor modifications. Both the ResNet50 model and squeezenet model could finish training and testing in a very short time (~8 min for ResNet50) and small epochs (20 epochs). In addition, these models successfully classified the cracked and no cracked images with accuracies of over 88% (ResNet50 model) and over 93% (squeezenet model).

For human tooth crack identification, we used an SSD model to locate the cracks. The COCO and VOC2007 datasets were applied as the training data. However, current datasets did not contain tooth crack images, so we collected 378 self-labeled crack images and added them to the dataset as a new class. The results demonstrated cracks could be recognized: one ground truth image was tested and recognized (Fig. 4a and 4b). Most of the ground truth crack was in the detect box; the upper part was out of the slot might due to the middle gap, which made the detector recognize them as two cracks since our detector could only detect one object. In addition, the cracks of new images of different sizes were also recognized, which indicated our tooth crack detector could assist the dentist in the crack diagnosis *in vitro*.

Modern imaging techniques, such as Ultrasound, MRI, and NIR imaging, play a vital role in disease diagnosis. However, the low-resolution images are considered a huge disadvantage of these methods, which may limit their applications. Super-resolution using the SRGAN model aims to improve the resolution and quality of the low-resolution image, which would be helpful for biomedical image analysis. Using the SRGAN model requires high-resolution images as the training dataset. In our tooth crack experiment, we trained the model with only COCO2014, the tooth crack remained the same without any enhancement because there were no high-resolution tooth crack images in the dataset. On the contrary, we used high-resolution concrete crack images as the training data because the resolution of our NIR tooth images is relatively low and the concrete crack image has the same main feature as the tooth crack image. The SR-enhanced tooth crack image became distinct and the crack turned dark from bright (Fig. 8d). When NIR light shone on the ICG-immersed human-extracted tooth from various directions, the crack could be bright instead of dark. The advantage of SRGAN

over SRResNet is the Discriminator, which could make the generated crack image closer to the original image.

Although modern imaging techniques have been well developed in recent decades, the resolution is still a limitation. More efforts need to be made for bio-medical image collection and classification, especially for high-resolution images. In addition, this work only detected the extracted human tooth. However, it would be more difficult to apply *in vivo* in clinical practice since teeth are next to each other. Our next step is to make a panoramic oral image/video using other machine-learning techniques that may solve the problem.

Conclusions

In this work, we build a human tooth crack dataset from NIR dental imaging videos for the first time. This dataset includes 593 cracked images and 601 non-cracked images; all the images are self-generated to a size of 400×600 by interpolation. Multiple deep learning-assisted models are innovatively implemented for tooth crack analysis. Two pre-trained classification models (ResNet50 model and squeezeNet1_1 model) had excellent classification performances, demonstrating that images with or without cracks can be well classified. Cracks are also recognized and located in the SSD model after training with the added tooth crack label in the COCO dataset. A high-resolution concrete crack image-based dataset can improve the resolution of tooth crack images by the SRGAN model. Our work represents the first effort to use machine-learning-aided NIRF dental imaging for automated dental disease identification, as compared to the previous studies of dental radiography. Overall, deep learning combined NIRF imaging system provides a novel tool for efficient and dependable human tooth crack diagnosis.

Acknowledgements The authors would like to extend their gratitude to the participants of the experiment. We want to thank Prof. Yanping Li (School of Environment and Sustainability, University of Saskatchewan) and Prof. Michelle L. Osborn (Department of Comparative Biomedical Science, School of Veterinary Medicine, Louisiana State University) for the suggestions and feedback on this work.

Funding This research was supported by NIH under Grant [U01AA029348], NSF CAREER award under Grant [2046929], and LSU Collaborative Cancer Research Initiative (010163).

Declarations

Conflict of interest All the authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

References

- Li, Z., et al. Mouthwash to deliver indocyanine green for near infrared dental fluorescence imaging. *IEEE J. Sel. Top. Quant. Electron.* 27(4):1–8, 2021. <https://doi.org/10.1109/JSTQE.2020.2995659>.
- Da Rosa, W. L., L. Brondani, T. Machado, E. Piva, and A. da Silva. Diagnosis and treatment of anterior cracked tooth: a case report. *Heighpubs J. Dent.* 2017. <https://doi.org/10.29328/journal.hjd.1001003>.
- Özcan, M., S. Abdin, and C. J. O. Sipahi. Bleaching induced tooth sensitivity: do the existing enamel craze lines increase sensitivity? A clinical study. *Odontology.* 102(2):197–202, 2014.
- Andreasen, J. O., S. Storgård Jensen, and V. Sae-Lim. The role of antibiotics in preventing healing complications after traumatic dental injuries: a literature review. *Endodont. Top.* 14(1):80–92, 2006. <https://doi.org/10.1111/j.1601-1546.2008.00231.x>.
- Hasan, S., K. Singh, and N. Salati. Cracked tooth syndrome: overview of literature. *Int. J. Appl. Basic Med. Res.* 5(3):164–168, 2015. <https://doi.org/10.4103/2229-516X.165376>.
- Pauwels, R. History of dental radiography: evolution of 2D and 3D imaging modalities. *Med. Phys. Int. J.* 8(1):235–277, 2020.
- Shadley, J. D., and S. J. M. Wolff. Very low doses of X-rays can cause human lymphocytes to become less susceptible to ionizing radiation. *Mutagenesis.* 2(2):95–96, 1987.
- Hsieh, Y., et al. Dental optical coherence tomography. *Sensors.* 13(7):8928–8949, 2013.
- Machoy, M., J. Seeliger, L. Szyszka-Sommerfeld, R. Koprowski, T. Gedrange, and K. Woźniak. The use of optical coherence tomography in dental diagnostics: a state-of-the-art review. *J. Healthc. Eng.* 2017(1):7560645, 2017.
- Stanga, P. E., E. Tsamis, A. Papayannis, F. Stringa, T. Cole, and A. Jalil. Swept-source optical coherence tomography Angio™ (Topcon Corp, Japan): technology review. *OCT Angiogr. Retinal Macular Dis.* 56:13–17, 2016.
- Kim, J., T. Shin, H. Kong, J. Y. Hwang, and H. Hyun. High-frequency ultrasound imaging for examination of early dental caries. *J. Dent. Res.* 98(3):363–367, 2019.
- Aljdaimi, A., H. Devlin, M. Dickinson, T. Burnett, and T. J. Slater. Micron-scale crack propagation in laser-irradiated enamel and dentine studied with nano-CT. *Clin. Oral Invest.* 23:2279–2285, 2019.
- Shah, N., N. Bansal, and A. Logani. Recent advances in imaging technologies in dentistry. *World J. Radiol.* 6(10):794–807, 2014. <https://doi.org/10.4329/wjr.v6.i10.794>.
- Li, Z., S. Yao, J. Xu, Y. Wu, C. Li, and Z. He. Endoscopic near-infrared dental imaging with indocyanine green: a pilot study. *Ann. N. Y. Acad. Sci.* 1421(1):88–96, 2018.
- Li, Z., et al. Optimal imaging windows of indocyanine green-assisted near-infrared dental imaging with rat model and its comparison to X-ray imaging. *J. Biophoton.* 13(6):e201960232, 2020.
- Li, Z., Z. Li, Y. Yang, S. Yao, C. Liu, and J. Xu. Original and liposome-modified indocyanine green-assisted fluorescence study with animal models. *Lasers Med. Sci.* 38(1):140, 2023. <https://doi.org/10.1007/s10103-023-03802-5>.
- Li, Z., et al. Indocyanine green-assisted dental imaging in the first and second near-infrared windows as compared with X-ray imaging. *Ann. N. Y. Acad. Sci.* 1448(1):42–51, 2019.
- Zhao, J., D. Zhong, and S. Zhou. NIR-I-to-NIR-II fluorescent nanomaterials for biomedical imaging and cancer therapy. *J. Mater. Chem. B.* 6(3):349–365, 2018.
- Li, Z., S. Yao, and J. Xu. "Indocyanine-green-assisted near-infrared dental imaging-the feasibility of in vivo imaging and the optimization of imaging conditions. *Sci. Rep.* 9(1):1–8, 2019.

20. Li, Z., et al. Mouthwash as a non-invasive method of indocyanine green delivery for near-infrared fluorescence dental imaging. *J. Biomed. Opt.* 27(6):066001, 2022.
21. Yu, M., J. Li, S. Liu, Z. Xie, J. Liu, and Y. Liu. Diagnosis of cracked tooth: clinical status and research progress. *Jpn. Dent. Sci. Rev.* 58:357–364, 2022.
22. Li, Z., et al. Detection and analysis of enamel cracks by ICG-NIR fluorescence dental imaging. *Ann. N. Y. Acad. Sci.* 1475(1):52–63, 2020.
23. Kim, J., S. Kang, and W. Yi. Automatic detection of tooth cracks in optical coherence tomography images. *J. Periodont. Implant Sci.* 47(1):41–50, 2017.
24. Aslam, M., et al. A novel method for detection of pancreatic ductal adenocarcinoma using explainable machine learning. *Comput. Methods Programs Biomed.* 245:108019, 2024. <https://doi.org/10.1016/j.cmpb.2024.108019>.
25. Kist, A. M., and M. Döllinger. Efficient biomedical image segmentation on EdgeTPUs at point of care. *IEEE Access.* 8:139356–139366, 2020.
26. Brattain, L. J., B. A. Telfer, M. Dhyani, J. R. Grajo, and A. E. Samir. Machine learning for medical ultrasound: status, methods, and future opportunities. *Abdominal Radiol.* 43(4):786–799, 2018.
27. Bloice, M. D., P. M. Roth, and A. Holzinger. Biomedical image augmentation using augmentor. *Bioinformatics.* 35(21):4522–4524, 2019.
28. de Haan, K., Y. Rivenson, Y. Wu, and A. Ozcan. Deep-learning-based image reconstruction and enhancement in optical microscopy. *Proc. IEEE.* 108(1):30–50, 2019.
29. Wang, Y., and J. Yi. Deep learning-based image registration method: with application to Scanning Laser Ophthalmoscopy (SLO) longitudinal images. In *Medical Imaging 2023: Image Processing*, vol. 12464. SPIE, 2023, pp. 601–605.
30. Li, Z., et al. Machine-learning-assisted spontaneous Raman spectroscopy classification and feature extraction for the diagnosis of human laryngeal cancer. *Comput. Biol. Med.* 146:105617, 2022.
31. Weigert, M., U. Schmidt, R. Haase, K. Sugawara, and G. Myers. Star-convex polyhedra for 3d object detection and segmentation in microscopy. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2020, pp. 3666–3673.
32. Elakkiya, R., K. S. S. Teja, L. Jegatha Deborah, C. Bisogni, and C. Medaglia. Imaging based cervical cancer diagnostics using small object detection-generative adversarial networks. *Multimedia Tools Appl.* 81(1):191–207, 2022.
33. Rashid, F., A. Irtaza, N. Nida, A. Javed, H. Malik, and K. M. Malik. Segmenting melanoma lesion using single shot detector (SSD) and level set segmentation technique. In *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*. IEEE, 2019, pp. 1–5.
34. Li, Y., B. Sixou, and F. Peyrin. A review of the deep learning methods for medical images super resolution problems. *IRBM.* 42(2):120–133, 2021.
35. Chen, Y., Y. Xie, Z. Zhou, F. Shi, A. G. Christodoulou, and D. Li. Brain MRI super resolution using 3D deep densely connected neural networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018, pp. 739–742.
36. Lee, S., S. Oh, J. Jo, S. Kang, Y. Shin, and J. Park. Deep learning for early dental caries detection in bitewing radiographs. *Sci. Rep.* 11(1):1–8, 2021.
37. Lee, J., D. Kim, S. Jeong, and S. Choi. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J. Dent.* 77:106–111, 2018.
38. Jaskari, J., et al. Deep learning method for mandibular canal segmentation in dental cone beam computed tomography volumes. *Sci. Rep.* 10(1):1–8, 2020.
39. Li, Z., et al. Cover image. *Ann. N. Y. Acad. Sci.* 1448:i–i, 2019. <https://doi.org/10.1111/nyas.14204>.
40. Li, Z., S. Yao, J. Xu, Y. Wu, C. Li, and Z. He. Cover image. *Ann. N. Y. Acad. Sci.* 1421(1):i, 2018. <https://doi.org/10.1111/nyas.13884>.
41. He, K., X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
42. Dwivedi, P. Detection of surface cracks in concrete structures using deep learning. <https://towardsdatascience.com/detection-of-surface-cracks-in-concrete-structures-using-deep-learning-f8f85cd8ac8b>.
43. Iandola, F. N., S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. arXiv preprint, 2016. arXiv:1602.07360.
44. Liu, W., et al. SSD: single shot multibox detector. In *European Conference on Computer Vision*, 2016, pp. 21–37.
45. Lin, T.-Y., et al. Microsoft coco: Common objects in context. In *European Conference on Computer Vision*, 2014, pp. 740–755.
46. Everingham, M., L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (VOC) challenge. *Int. J. Comput. Vis.* 88(2):303–338, 2010.
47. Tzatalin, D. LabelImg. GitHub Repository 6, 2015.
48. Ledig, C., et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4681–4690.
49. Özgenel, Ç. F., and A. G. Söğüt. Performance comparison of pretrained convolutional neural networks on crack detection in buildings. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, vol. 35. IAARC Publications, 2018, pp. 1–8.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.