

Defect Segmentation in Concrete Structures Combining Registered Infrared and Visible Images: A Comparative Experimental Study [†]

Sandra Pozzer ^{1,*} , Marcos Paulo Vieira de Souza ¹, Bata Hena ¹, Reza Khoshkbary Rezaiye ¹, Setayesh Hesam ¹, Fernando Lopez ² and Xavier Maldague ¹ 

¹ Department of Electrical and Computer Engineering, Laval University, 1065, Av., de la Médecine, Québec City, QC G1V 0A6, Canada; marcos-paulo.vieira-de-souza.1@ulaval.ca (M.P.V.d.S.); bata-nkirda.hena.1@ulaval.ca (B.H.); reza.khoshkbary-rezaiye.1@ulaval.ca (R.K.R.); setayesh.hesam.1@ulaval.ca (S.H.); xavier.maldague@gel.ulaval.ca (X.M.)

² TORNGATS, Québec City, QC G2E 5V9, Canada; flopez@torngats.ca

* Correspondence: sandra.pozzer.1@ulaval.ca

[†] Presented at the 16th International Workshop on Advanced Infrared Technology & Applications, 26–28 October 2021; Available online: <https://aita2021.sciforum.net/>.

Abstract: This study investigates the semantic segmentation of common concrete defects when using different imaging modalities. One pre-trained Convolutional Neural Network (CNN) model was trained via transfer learning and tested to detect concrete defect indications, such as cracks, spalling, and internal voids. The model's performance was compared using datasets of visible, thermal, and fused images. The data were collected from four different concrete structures and built using four infrared cameras that have different sensitivities and resolutions, with imaging campaigns conducted during autumn, summer, and winter periods. Although specific defects can be detected in monomodal images, the results demonstrate that a larger number of defect classes can be accurately detected using multimodal fused images with the same viewpoint and resolution of the single-sensor image.

Keywords: non-destructive testing; deep learning; image processing; multimodal images; infrared thermography; concrete bridges; infrastructure inspection



Citation: Pozzer, S.; Souza, M.P.V.d.; Hena, B.; Rezaiye, R.K.; Hesam, S.; Lopez, F.; Maldague, X. Defect Segmentation in Concrete Structures Combining Registered Infrared and Visible Images: A Comparative Experimental Study. *Eng. Proc.* **2021**, *8*, 29. <https://doi.org/10.3390/engproc2021008029>

Academic Editors: Giovanni Ferrarini, Paolo Bison and Gianluca Cadelano

Published: 1 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Computer vision methods combined with remote sensors represent a promising solution for the automatic inspection of concrete structures [1,2]. The goal of using such advanced non-destructive testing and evaluation (NDTE) methods is to more efficiently evaluate the inspection of large infrastructures [3]. The defects present in concrete structures such as cracks, spalling, and internal voids, can appear with a particular color, shape, and thermal contrast in the images acquired in the visible and/or infrared spectrum. For instance, the detection of subsurface voids is not possible using visible images only, and the occurrence of spalling and cracks are not sufficiently clear in thermal images. This scenario represents a challenge for human and computer analyses, where multimodal images from the same structure and field of view must be inspected multiple times due to their diverse information content.

Given this outline, the development of registering and fusion techniques to integrate with multimodal images has been studied and receiving attention [4]. As a result, the application of smart algorithms for automated feature detection on fused images is becoming more common [5,6], as the multimodal image can intelligently combine the complementary information with the several input sensors. However, there is a valid concern about whether the pixel fusion of the multiple input images can cause a loss of information for feature detection [7]. In this study, the segmentation performance of concrete defect

indications commonly present in visible and thermal images is investigated and compared using multimodal and fused images. The hypothesis is that the registration and fusion of the multiple-sensor images can bring information to the learning algorithm without a significant loss of information. Moreover, the merged images can speed up the labeling, training, and evaluation of the asset by using only one dataset of fused images.

2. Methods

The data acquisition campaigns were performed in 4 concrete structures: a concrete dam, 2 highway bridges, and 1 railway overpass in Canada. Part of the current dataset was collected and analyzed in a previous study [2]. Four different thermal cameras were used to collect the data, and all the devices were equipped with both thermal and optical sensors; therefore, each capture provides an optical and thermal frame with the same field of view, allowing further image registration and fusion. The imagery campaigns were conducted at different times of the year to obtain different thermal information. All the experimental tests were performed under passive heating.

A total of 690 visible and infrared images resulted from the inspection campaigns. The further applied computer-based procedures can be summarized: (1) The correlated single-sensors visible and infrared images were downloaded and structured inside 2 datasets. (2) All images were registered and fused. (3) Original visible and infrared images were replaced by their respective registered images. (4) A dataset of fused images was created, with images of the same resolution and from the same field of view of the registered infrared and visible images. (5) All cracks and spalling were labeled using the visible images dataset, and the internal voids indications were labeled using the thermal image dataset. (6) Each dataset was split into training (85%), validation, and testing datasets (15%). (7) Different image processing techniques were applied to the datasets: contrast enhancement using histogram equalization and resolution improvement using the Very Deep Convolutional Networks (VDCN) algorithm [8] trained in visible and thermal images. (8) A pre-trained Convolutional Neural Network (CNN) model, i.e., MobileNetV2 [9], was trained (via transfer learning), validated, and tested to detect the multiclass labeled defects using the different datasets. (9) The segmentation performance was evaluated in each scenario using the Intersection over Union, Pixel Accuracy, and Mean F1 Score metrics. Given the different acquisition scenarios, all data were normalized before the analysis.

3. Results

The CNN model demonstrated the possibility of accurately identifying the investigated damages in the multimodal and fused images of concrete structures. Figure 1 shows one example of the tested images that contains crack and spalling. The segmentation results are shown according to the different image modalities and image processing techniques.

The VDCN algorithms visually improved the image resolution in all cases, proving to be suitable for use in thermal, visible, and fused images. The visual result of the segmentation shows that the model was able to identify most of the defects and that the image processing techniques can help the model be more specific and avoid false positives, as we can see in the comparison of Figure 1c with Figure 1h,m,t. The results of IOU, Pixel Accuracy, and Mean F1 Score were also computed and will be provided later on. The outcome showed that image processing techniques can increase defect detection. However, in some cases, the interventions led to a decrease in the segmentation performance, such as in the case of crack detection. When it comes to multiple damage detection, some filters can overpass or hide damages that have individual characteristics (color, shape, texture, etc.). By way of illustration, contrast enhancement can facilitate spalling identification but can cause thin cracks to vanish or change the appearance and texture of the concrete portions. In addition, both thermal and visible images contain noise that are consequently combined when the images are fused. The use of specific image processing techniques to reduce the level of noise before and after the fusion and segmentation may improve the model performance. In this experiment, the aim was to compare the segmentation performance

using different image datasets (multimodal and fused images), and further research will be performed to improve the deep learning model to pursue a better performance in multiclass segmentation.

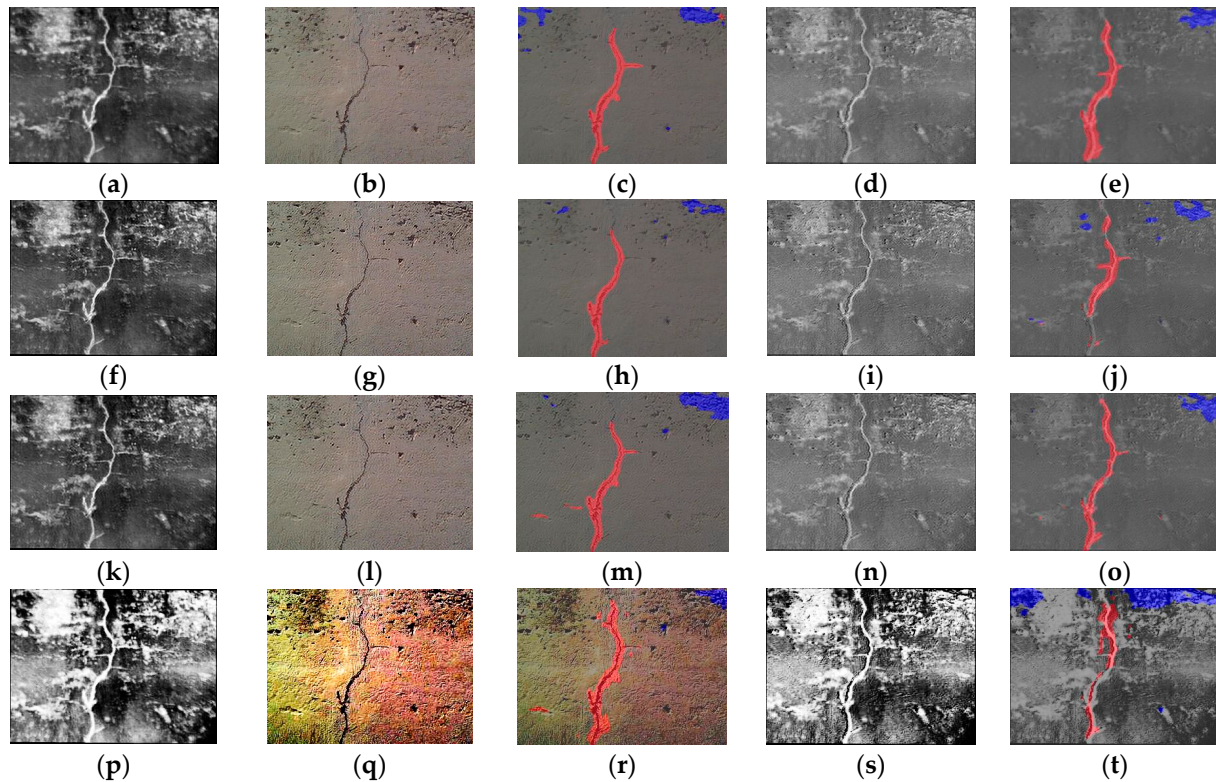


Figure 1. Visual results of image processing techniques and multiclass damage detection on the multimodal images, where red represents the predicted cracks and blue represents the predicted spallings in each case; the columns show thermal images, visible images, segmentation results on visible images, visible and infrared fused images, and segmentation results on fused images, respectively (left to right). The first row (a–e) shows examples of the original dataset; the second row (f–j) shows examples of the improved dataset based on the Very Deep Convolutional Networks algorithm trained in visible images; The third row (k–o) shows examples of the improved dataset based on the Very Deep Convolutional Networks algorithm trained in thermal images; and the last row (p–t) shows examples of the dataset with the contrast enhanced by image histogram equalization.

In a defect class analysis, the models yield rather similar results in segmentation for processed multimodal and fused images from the same field of view. A hypothesis test was conducted using a two-way analysis of variance (ANOVA) to verify if there was any significant difference in the segmentation performance results when using the different image modalities and fused images. Considering a level of significance of 0.01, the results show that statistically, there is no significant difference between the average IOU, accuracy, and F1score when the different image datasets are used as input of the CNN model. This preliminary result indicates that fused images, with the same field of view and resolution as the single-sensor image, can be used for the computation of cracks, internal voids, and spalling in concrete structures without any significant harm to the accuracy of the defect segmentation. The outcome is advantageous for the computer vision and analyst point of view, as the fused images can contain a higher number of defect classes than the singular images. Therefore, the use of fused images can save the time, effort, and resources spent in the labeling and training of multimodal images, as they smartly use the combination complementary information of multiple input sensors in one single image.

Future research could use a larger number of thermal, visible, and fused images from other types of concrete structures or use diverse NDT sensors, as well as explore the

use of automated labeling for one or multiclass damages in different images modalities. Another possible area of future research would be to investigate the use of two multimodal images as the input of the network, exploring the characteristics of feature detection in each modality before the fusion and segmentation, as well as modifying the model to add a deep feature extraction layer, for instance. In addition, the effect of using different fusion algorithms on the performance of CNN could be investigated.

Regarding the data availability, the datasets, model, and codes that support the findings of this study are available from the corresponding author upon reasonable request.

4. Conclusions

A comparative experimental study was performed to investigate crack, spalling, and internal void segmentation in concrete structures using datasets of visible, thermal, and fused images with the same resolution and point of view. The experiment was developed using CNN and image processing techniques, which automatically identified concrete damages in the multimodal and fused images from different concrete structures. The contribution of this study is attributed to the (1) automatic defect detection improvement based on the fusion of infrared and visible images due to higher information content, (2) the possibility of defect segmentation improvement based on the application of simple image processing techniques, such as image resolution improvement and image contrast enhancement, (3) automated crack, spalling, and delamination segmentation with a multi-class convolutional neural network, and (4) use in the growing field of computer vision applied to concrete bridge inspection.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Spencer, B.F.; Hoskere, V.; Narazaki, Y. Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. *Engineering* **2019**, *5*, 199–222. [\[CrossRef\]](#)
2. Pozzer, S.; Rezazadeh Azar, E.; Dalla Rosa, F.; Chamberlain Pravia, Z.M. Semantic Segmentation of Defects in Infrared Thermographic Images of Highly Damaged Concrete Structures. *J. Perform. Constr. Facil.* **2021**, *35*, 04020131. [\[CrossRef\]](#)
3. Koch, C.; Paal, S.G.; Rashidi, A.; Zhu, Z.; König, M.; Brilakis, I. Achievements and Challenges in Machine Vision-Based Inspection of Large Concrete Structures. *Adv. Struct. Eng.* **2014**, *17*, 303–318. [\[CrossRef\]](#)
4. Ma, J.; Liang, P.; Yu, W.; Chen, C.; Guo, X.; Wu, J.; Jiang, J. Infrared and Visible Image Fusion via Detail Preserving Adversarial Learning. *Inf. Fusion* **2020**, *54*, 85–98. [\[CrossRef\]](#)
5. Yousefi, B.; Sfarra, S.; Ibarra-Castanedo, C.; Avdelidis, N.P.; Maldague, X.P.V. Thermography Data Fusion and Nonnegative Matrix Factorization for the Evaluation of Cultural Heritage Objects and Buildings. *J. Therm. Anal. Calorim.* **2019**, *136*, 943–955. [\[CrossRef\]](#)
6. Lim, H.J.; Hwang, S.; Kim, H.; Sohn, H. Steel Bridge Corrosion Inspection with Combined Vision and Thermographic Images. *Struct. Health Monit.* **2021**, *20*, 1475921721989407. [\[CrossRef\]](#)
7. Xydeas, C.; Petrović, V. Pixel-level image fusion metrics. In *Image Fusion*; Elsevier: Amsterdam, The Netherlands, 2008; pp. 429–450. ISBN 978-0-12-372529-5.
8. Kim, J.; Lee, J.K.; Lee, K.M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. *arXiv* **2016**, arXiv:1511.04587.
9. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. *arXiv* **2019**, arXiv:1801.04381.