

Proceeding Paper

# Building Façade Protection Using Spatial and Temporal Deep Learning Models Applied to Thermographic Data. Laboratory Tests <sup>†</sup>

Iván Garrido <sup>1,\*</sup> , Eva Barreira <sup>2</sup> , Ricardo M. S. F. Almeida <sup>2,3</sup> and Susana Lagüela <sup>4</sup> 

<sup>1</sup> GeoTECH Group, CINTECX, Universidade de Vigo, Campus Universitario de Vigo, As Lagoas, Marcosende, 36310 Vigo, Spain

<sup>2</sup> CONSTRUCT-LFC, Civil Engineering Department, Faculty of Engineering (FEUP), University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal; barreira@fe.up.pt (E.B.); ralmeida@estgv.ipv.pt (R.M.S.F.A.)

<sup>3</sup> Department of Civil Engineering, Polytechnic Institute of Viseu, Campus Politécnico, 3504-510 Viseu, Portugal

<sup>4</sup> Department of Cartographic and Terrain Engineering, EPS Ávila, University of Salamanca, Calle Hornos Caleros 50, 05003 Ávila, Spain; sulaguela@usal.es

\* Correspondence: ivgarrido@uvigo.es

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

**Abstract:** This paper proposes a methodology that combines spatial and temporal deep learning (DL) models applied to data acquired by InfraRed Thermography (IRT). The data were acquired from laboratory specimens that simulate building façades. The spatial DL model (Mask Region-Convolution Neural Network, Mask R-CNN) is used to identify and classify different artificial subsurface defects, whereas the temporal DL model (Gated Recurrent Unit, GRU) is utilized to estimate the depth of each defect, all in an autonomous and automated manner. An F-score average of  $92.8 \pm 5.4\%$  regarding defect identification and classification, and a root-mean-square error equal to 1 mm in the estimation of defect depth equal to 10 mm as the best defect depth estimation, are obtained with this first application of a combination of spatial and temporal DL models to the IRT inspection of buildings.

**Keywords:** InfraRed Thermography; deep learning; defect; identification; classification; depth estimation; laboratory tests; building façade



check for  
updates

**Citation:** Garrido, I.; Barreira, E.; Almeida, R.M.S.F.; Lagüela, S. Building Façade Protection Using Spatial and Temporal Deep Learning Models Applied to Thermographic Data. Laboratory Tests. *Eng. Proc.* **2021**, *8*, 20. <https://doi.org/10.3390/engproc2021008020>

Academic Editors: Giovanni Ferrarini, Paolo Bison and Gianluca Cadelano

Published: 23 November 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

Preventive and conservation interventions in buildings are essential tasks to protect the physical assets that provide habitat and comfort to people against deterioration caused by the passage of time. This requires the use of the most advanced inspection technologies and their most advanced data processing algorithms to identify and characterize building defects, from those in the early stages of growth to the most severe ones.

InfraRed Thermography (IRT), in addition to the non-invasive nature common to all non-destructive techniques, has the advantage of being a non-contact inspection technique that allows: (i) operating in real time, (ii) scanning at high speed, and (iii) inspecting a wide area coverage of the object under study, in this case, the building envelope [1]. However, there are some drawbacks found in traditional IRT data processing algorithms that still need to be solved, such as: (i) the low resolution of infrared cameras, (ii) the high dependence on environmental conditions, (iii) the homogeneous heating and/or cooling that needs to be provided on the surface under study, and (iv) the control of the different mechanisms of heat transfer.

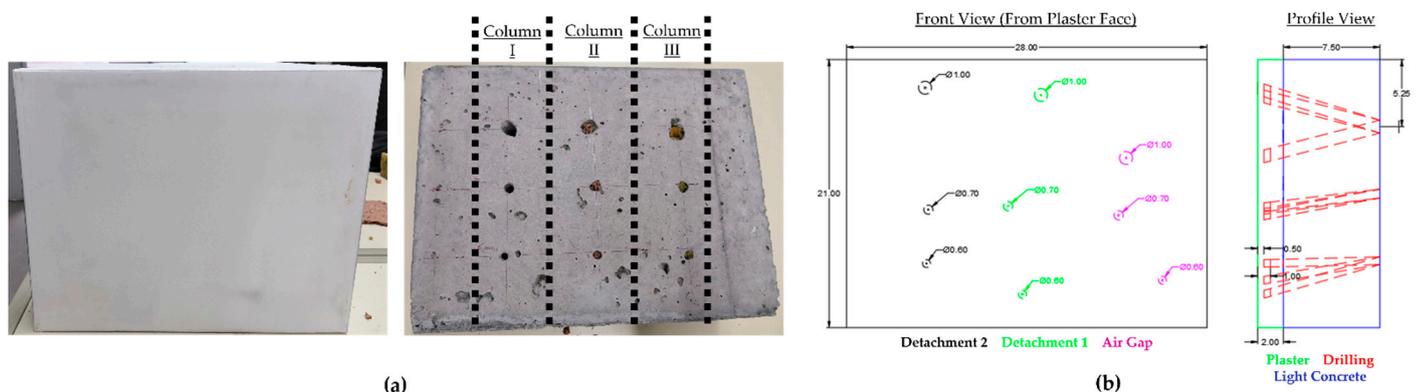
In this situation, deep learning (DL) models are recent data processors that have proven to be robust to highly complex problems, solving them in an autonomous and automatic manner. The positive results offered by different DL models have led to their use

and adaptation to the IRT field, addressing the aforementioned drawbacks of the traditional algorithms. Garrido et al. [2] and Al-Habaibeh et al. [3] are examples of the most recent papers applying DL models to thermographic inspection of buildings. However, either only a qualitative analysis (identification and classification of water-related problems and thermal bridges [2]) or only a quantitative analysis (prediction of future savings after a retrofitting [3]) is performed, an indicator that still shows the scarcity of IRT papers using DL in buildings.

With the aim of widening the application of DL models to thermographic inspection of buildings, this paper uses a spatial and a temporal DL model applied to thermographic sequences acquired from three laboratory specimens that simulate a building façade to: (i) automatically identify and classify different types of artificial subsurface defects (air gaps and two types of detachment) with the spatial DL model (Mask Region-Convolution Neural Network, Mask R-CNN), and (ii) automatically estimate the depth of each of them with the temporal DL model (Gated Recurrent Unit, GRU).

## 2. Materials and Methods

The three laboratory specimens consist of plaster-coated lightweight concrete with artificial subsurface defects simulating air gaps and two different types of detachment. The dimensions of the specimen and defects, as well as the spatial distribution of the latter, are the same for the three specimens. The difference between them is the depth of the subsurface defects with respect to the plaster surface: 0.5 cm (specimen 1), 1 cm (specimen 2) and 2 cm (specimen 3). Figure 1 shows visible images of one of the specimens (Figure 1a) and their outline (Figure 1b).

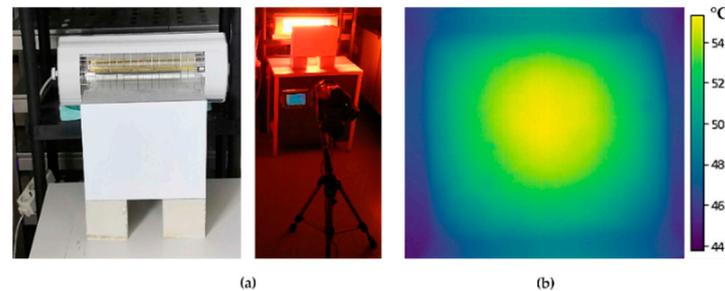


**Figure 1.** (a) Visible images of one of the specimens. (b) Outline of the specimens (units in cm).

Figure 1b shows that each defect is created by a perforation with a specific diameter (between 0.6, 0.7 and 1 cm) and inclination from the opposite face of the plaster surface. In total there are 9 perforations in each specimen, leaving three hollow perforations (simulating the air gaps), the other three introducing crushed cork (simulating one of the detachment types), and the last three introducing crushed sponge (simulating the other detachment type). Each type of defect is distributed in one of the three different columns (columns I, II and III) and with the three diameters of perforation in each column, as shown in Figure 1b.

Figure 2a represents the experimental setup, which is the same for each specimen. An IR camera (NEC Avio TH9100MR, 320 (Horizontal)  $\times$  240 (Vertical) pixels, 8  $\mu$ m to 14  $\mu$ m) is placed in front of the plaster surface, at 70 cm distance and with a horizontal deviation angle of about 5° in the IR camera position with respect to the specimen plane to avoid any possible reflection on the surface. A 2-kW infrared lamp is placed centered at a distance of 20 cm from the lightweight concrete face. Using the lamp, a one-hour heating is performed followed by a 30 min non-heating while thermal images are acquired at the rate of one

image every 30 s, in order to capture any possible growth of the thermal footprint of the defects (Figure 2b). This test is performed for each specimen, both with the orientation shown in Figure 1a,b and by rotating 180° to increase the acquired dataset.



**Figure 2.** (a) Experimental setup. (b) One of the thermal images acquired.

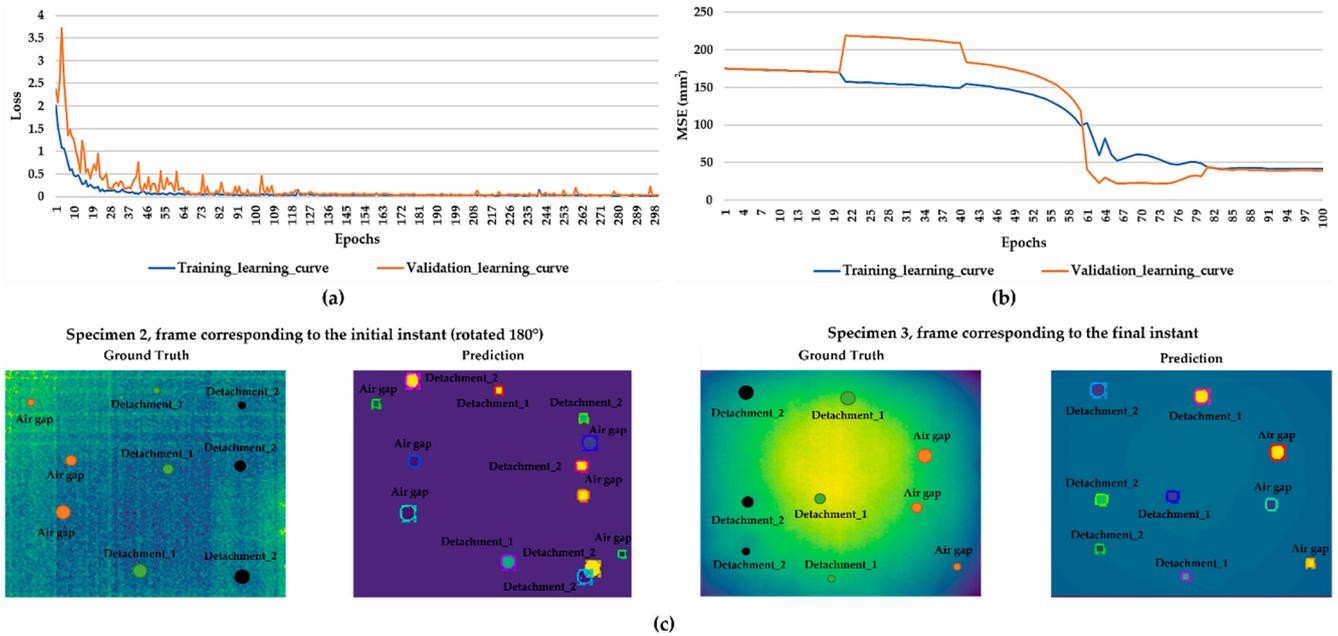
### 3. Results and Discussion

Each thermal image sequence is processed from the beginning of the appearance of the thermal footprints of the subsurface defects to the maximum representation of these footprints in each test. In total, 826 thermal images are of interest after the tests performed on the three specimens, between 132 and 141 thermal images of interest in each test.

So far, Mask R-CNN is one of the best spatial DL models, which was developed by the Facebook Artificial Intelligence Research [4]. For the Mask R-CNN training process, the dataset is first shuffled, and the k-fold cross validation and 90-10 train-test split approach are used, having approximately the same proportion with respect to the number of images of each test in the testing dataset. Five folds are selected for the original training dataset, thus having the following percentage values: (i) 72% training dataset, (ii) 18% validation dataset, and (iii) 10% testing dataset. In addition, (i) transfer learning is applied, (ii) all the layers of Mask R-CNN are updated at each epoch, (iii) batch gradient descent is selected, (iv) a tensor shape of  $(512 \times 512 \times 3)$  is applied to each input element of the dataset, and (v) the value set for epoch, learning rate, learning momentum and weight decay is 300 (60 for each fold), 0.001, 0.9, and 0.1, respectively.

Moreover, GRU was proposed by Cho et al. [5], and has been successfully applied to sequential or temporal data. The input dataset to the GRU model are thermal contrast vectors extracted from the defect areas. The difference between the maximum temperature of a defect and its surrounding temperature at each instant of a test is calculated in a vector. Then, the thermal images of interest from each sequence are reshaped into vectors. A total of 54 thermal contrast vectors are obtained, applying a standardization process on each vector. A total of 9 vectors corresponding to different depths ( $\times 3$ ) and defect types ( $\times 3$ ), but with the same defect diameter (0.6 cm), are used as the testing dataset. In this way, the robustness of the DL model is measured at the lowest thermal contrasts, i.e., at the smallest defect areas. The remaining 45 vectors are shuffled and the 80-20 train-validate split approach is used, implementing k-fold cross validation with “k” equal to 5. Before the GRU training process, the following values are set for the neurons (units in the GRU cell), hidden layers (GRU cells), learning rate, batch size and epochs, respectively: 200, 2, 0.0001, batch gradient descent and 100 (20 for each fold). In addition, the Mean Square Error (MSE) function is used as the cost function, Rectified Linear Unit (ReLU) as the activation layer and Adaptive Moment Estimation (Adam) as the optimizer.

The source code of [6] is used for the Mask R-CNN implementation, using [7] for GRU. Figure 3a,b show the evolution of the training-validation loss and training-validation MSE during the Mask R-CNN and GRU training, respectively.



**Figure 3.** (a) Mask R-CNN training result. (b) GRU training result. (c) Mask R-CNN predictions on some images from the testing dataset.

Observing Figure 3a, the training and validation curves converge towards the end of the second fold (around the 109th epoch). The evolutions of the training and validation loss values are stabilized after the convergence point, around 0.033 and 0.053, respectively. As for Figure 3b, the training and validation curves converge and stabilize at the beginning of the last fold (around the 80th epoch). Training validation MSE is around 41 mm<sup>2</sup> from the convergence point (6.4 mm applying the root-mean-square Error (RMSE)). Tables 1 and 2 represent two confusion matrices after applying the trained Mask R-CNN to the corresponding testing dataset: one showing the ability of the DL model to correctly differentiate a pixel belonging to a defect and a non-defect (Table 1), and the other showing the ability of the DL model to correctly classify the defect type with respect to the true positives, i.e., the pixels correctly assigned as defects (Table 2). Note that the confidence score is set at 0.7. Table 3 represents the results of the defect depth estimation after applying the trained GRU to the corresponding testing dataset.

**Table 1.** Results after applying the trained Mask R-CNN to the testing dataset. Part I—defect identification.

		Actual Class	
		Defect	Non-Defect
Predicted Class	Defect	3264.4 ± 113.4	38 ± 283.7
	Non-Defect	349.3 ± 111.6	258,492.2 ± 294.8
Precision (%)		95.4 ± 4.6	
Recall (%)		90.3 ± 4.4	
F-score (%)		92.8 ± 5.4	

Tables 1 and 2 show that a high number of true positives and true negatives are obtained. The precision, recall, and F-score results are also shown in Table 1, confirming these good results. It can be seen that the trained Mask R-CNN is more robust to false positives than to false negatives, obtaining the worst values in the images corresponding to the initial instants (see Figure 3c). Table 2 further indicates a perfect overall accuracy, which means that the trained Mask R-CNN does not misclassify any true positive. As for Table 3,

only acceptable results are obtained in the estimation of defect depth equal to 10 mm (RMSE equal to 1 mm) and to 20 mm (RMSE equal to 7.1 mm), being, therefore, necessary to increase the dataset size before the GRU training. It is noted that the need for additional tests is also important to increase the robustness of the trained Mask R-CNN, especially considering tests in less controlled scenarios such as inspections on real building façades.

**Table 2.** Results after applying the trained Mask R-CNN to the testing dataset. Part II—defect classification.

		Actual Class		
		Air Gap	Detachment 1	Detachment 2
Predicted Class	Air Gap	1098.6 ± 30.9	0	0
	Detachment 1	0	1081.5 ± 101.6	0
	Detachment 2	0	0	1084.3 ± 16.9
<b>Overall Accuracy (%)</b>		100		

**Table 3.** Results after applying the trained GRU to the testing dataset.

Expected Depth (mm)	Estimated Depth (mm)	RMSE (mm)
5	9.8 ± 1.4	4.9
10	10.8 ± 0.9	1
20	13 ± 0.8	7.1

#### 4. Conclusions

The contribution of this paper is the proposal of a methodology that advances the optimization of decision making in prevention and conservation actions in the building field. The combination of spatial and temporal DL models can be applied to real building façades, first identifying and classifying defects with the spatial DL model, and then estimating the depth of each identified defect area with the temporal DL model. Future research will continue to use spatial and temporal DL models in IRT imagery by applying them to other materials and also to other defects with the aim at increasing the robustness of DL models applied to thermographic inspections.

**Author Contributions:** Conceptualization, I.G., E.B., R.M.S.F.A. and S.L.; methodology, I.G.; software, I.G.; validation, I.G., E.B.; formal analysis, I.G., E.B., R.M.S.F.A. and S.L.; investigation, I.G. and E.B.; resources, E.B. and R.M.S.F.A.; data curation, I.G.; writing—original draft preparation, I.G.; writing—review and editing, I.G., E.B., R.M.S.F.A. and S.L.; visualization, I.G., E.B., R.M.S.F.A. and S.L.; supervision, E.B., R.M.S.F.A. and S.L.; funding acquisition, I.G., E.B. and R.M.S.F.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** Authors would like to thank the Ministerio de Ciencia, Innovación y Universidades (Gobierno de España) for the financial support given through programs for human resources (FPU16/03950) and the Programa IACOBUS 2021 for the financial support given through programs for short stay mobilities (candidacy number: 99).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This work has used computational resources from the Supercomputing Centre of Galicia (CESGA).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Garrido, I.; Lagüela, S.; Otero, R.; Arias, P. Thermographic methodologies used in infrastructure inspection: A review—Post-processing procedures. *Appl. Energy* **2020**, *266*, 114857. [[CrossRef](#)]
2. Garrido, I.; Lagüela, S.; Fang, Q.; Arias, P. Combination of thermal fundamentals and Deep Learning for infrastructure inspections from thermographic images. Preliminary results. In Proceedings of the 15th Quantitative InfraRed Thermography, Porto, Portugal, 6–10 July 2020.
3. Al-Habaibeh, A.; Sen, A.; Chilton, J. Evaluation Tool for the Thermal Performance of Retrofitted Buildings Using An Integrated Approach of Deep Learning Artificial Neural Networks and Infrared Thermography. *Energy Built Environ.* **2021**, *2*, 345–365. [[CrossRef](#)]
4. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 386–397. [[CrossRef](#)] [[PubMed](#)]
5. Cho, K.; van Merriënboer, B.; Bahdanau, D.; Bengio, Y. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. *arXiv* **2015**, arXiv:1409.1259. [[CrossRef](#)]
6. Mask\_RCNN/Samples/Balloon at Master Matterport/Mask\_RCNN GitHub. Available online: [https://github.com/matterport/Mask\\_RCNN/tree/master/samples/balloon](https://github.com/matterport/Mask_RCNN/tree/master/samples/balloon) (accessed on 27 July 2021).
7. GitHub-Umeshpalai/AlgorithmicTrading-MachineLearning: Deep Learning-Neural network (RNN, LSTM & GRU). Available online: <https://github.com/umeshpalai/AlgorithmicTrading-MachineLearning> (accessed on 5 August 2021).