Detection of exposed steel rebars based on deep-learning techniques and unmanned aerial vehicles

R. Santos [a], D. Ribeiro [a,*], P. Lopes [b], R. Cabral [c], R. Calçada [c]

[a] CONSTRUCT - LESE, School of Engineering, Polytechnic of Porto, Porto, Portugal
[b] School of Engineering, Polytechnic of Porto, Porto, Portugal
[c] CONSTRUCT - LESE, Faculty of Engineering, University of Porto, Porto, Portugal

ABSTRACT

In recent years deep-learning techniques have been developed and applied to inspect cracks in RC structures. The accuracy of these techniques leads to believe that they may also be applied to the identification of other pathologies. This article proposes a technique for automated detection of exposed steel rebars. The tools developed rely on convolutional neural networks (CNNs) based on transfer-learning using AlexNet. Experiments were conducted in large-scale structures to assess the efficiency of the method. To circumvent limitations on the proximity access to structures as large as the ones used in the experiments, as well as increase cost efficiency, the image capture was performed using an unmanned aerial system (UAS). The final goal of the proposed methodology is to generate orthomosaic maps of the pathologies or structure 3D models with superimposed pathologies. The results obtained are promising, confirming the high adaptability of CNN based methodologies for structural inspection.

1. Introduction

Reinforced concrete is one of the most common materials used worldwide on the construction of civil engineering structures. Its widespread utilization is a consequence of the combination of several important factors such as, among others, low to medium complexity of execution, workmanship expertise for common low-rise buildings, general durability and the ability of the material to be easily cast on the desired shapes [1]. In time, however, reinforced concrete structures inevitably begin to suffer from anomalies that compromise their durability and functionality.

Generally, the main anomalies detected in reinforced concrete are cracks [2–4], efflorescences [5], corrosion of steel elements [6,7], delamination [8,9] and exposed steel rebar reinforcements [10,11]. Among these anomalies, the corrosion of rebar reinforcements is the main cause of damage and early failure of reinforced concrete structures [12]. Consequently, corrosion of rebar reinforcements must be detected through continuous monitoring of civil engineering structures and treated as early as possible in order to ensure their integrity.

The corrosion of rebar reinforcements is typically examined by non-destructive evaluation technologies. Current bibliography and recent research mention several different studies on the detection of concrete

reinforcements corrosion, with techniques based both on early stages without concrete delamination or exposed steel rebars and on advanced stages with concrete delamination accompanied by exposed steel rebars.

The methods used for non-destructive evaluation of early-stage rebar reinforcement corrosion are mainly based on electromagnetic sensors. Devices emit an electromagnetic pulse and detect the magnetic field induced in metal objects allowing the determination of the location where the embedded metal becomes exposed. The most common techniques involve the use of ultrasonic waves [13,14], Ground Penetrating Radar (GPR) [15,16], InfraRed Thermography (IRT) [17], gamma-ray radiography [18] or other electromagnetic sensors [19].

On advanced stages of deterioration visual inspection is the method normally used for detection, examination and assessment of the evolution of concrete structures pathologies. For visual inspection reports, images can be taken by handheld cameras or Unmanned Aerial Vehicles (UAVs). Inspectors visually evaluate defects based on experience, skill, and engineering discernment. Nonetheless, this process is subjective, laborious, time-consuming and hampered by the need for closed proximity access to several parts of complex structures.

Researchers and engineers have been proposing automated visual inspection methods based on recent developments on digital image processing techniques [20–23] and on artificial intelligence techniques
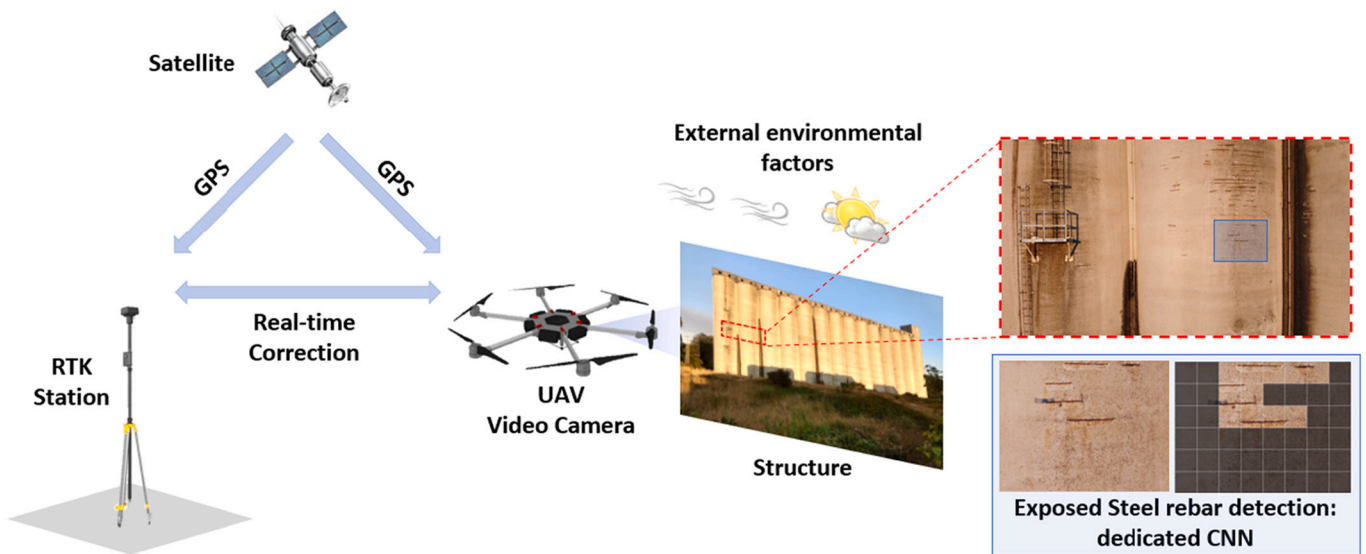

Fig. 1. General system description.

Table 1
Descriptions of the system components.

BRAND/MODEL		TECHNICAL CHARACTERISTICS
DJI Mavic Mini		Takeoff Weight: 249 g Max Flight Time: Approx. 30 minutes Hovering Accuracy Range: Vert.: ±0.5 m Horiz.: ±1.5 m
		Camera: Integrated Sensor: 1/2.3" CMOS Effective Pixels: 12 MP Image Size: 4000×2250p Photo Format: JPG
DJI Matrice 600 Pro		Max Takeoff Weight: 9.5 kg + 6 kg (payload) Max Flight Time: Approx. 32 minutes Hovering Accuracy Range: Vert.: ±0.5 m Horiz.: ±1.5 m
		Camera: DJI Zenmuse X5 Sensor: 4/3" CMOS Effective Pixels: 16 MP Image Size: 4608×3456p Photo Format: JPG

[24–28], in order to ensure safer, more efficient and economical structural assessment [29].

Digital image processing techniques deal with the problems of applying image filters in order to highlight surface non-uniformities. Nevertheless, the analysis of images with strong light or shadow casting present difficult or unsurmountable challenges that may hinder the application of these techniques. Furthermore, it is not generally an autonomous method, and the parameters of the several image filters that may be applied must be adjusted by the inspector.

In contrast to digital image processing techniques, artificial intelligence techniques allow autonomous identification of surface anomalies in the image, typically performed using deep learning. Deep learning has been particularly successful in surpassing human capabilities in visual image recognition, speech recognition and natural language processing. Convolutional Neural Networks (CNNs) are a class of deep neural networks sketched for processing structured arrays of data such as images, that have revolutionized autonomous image classification and object identification in the last decade.
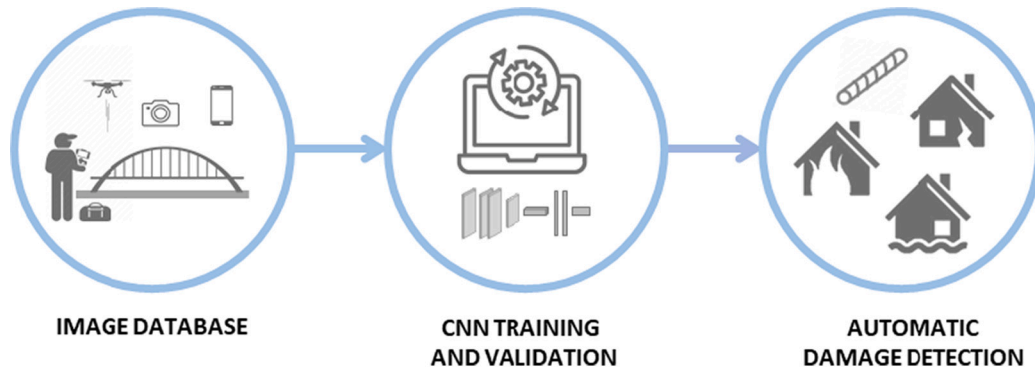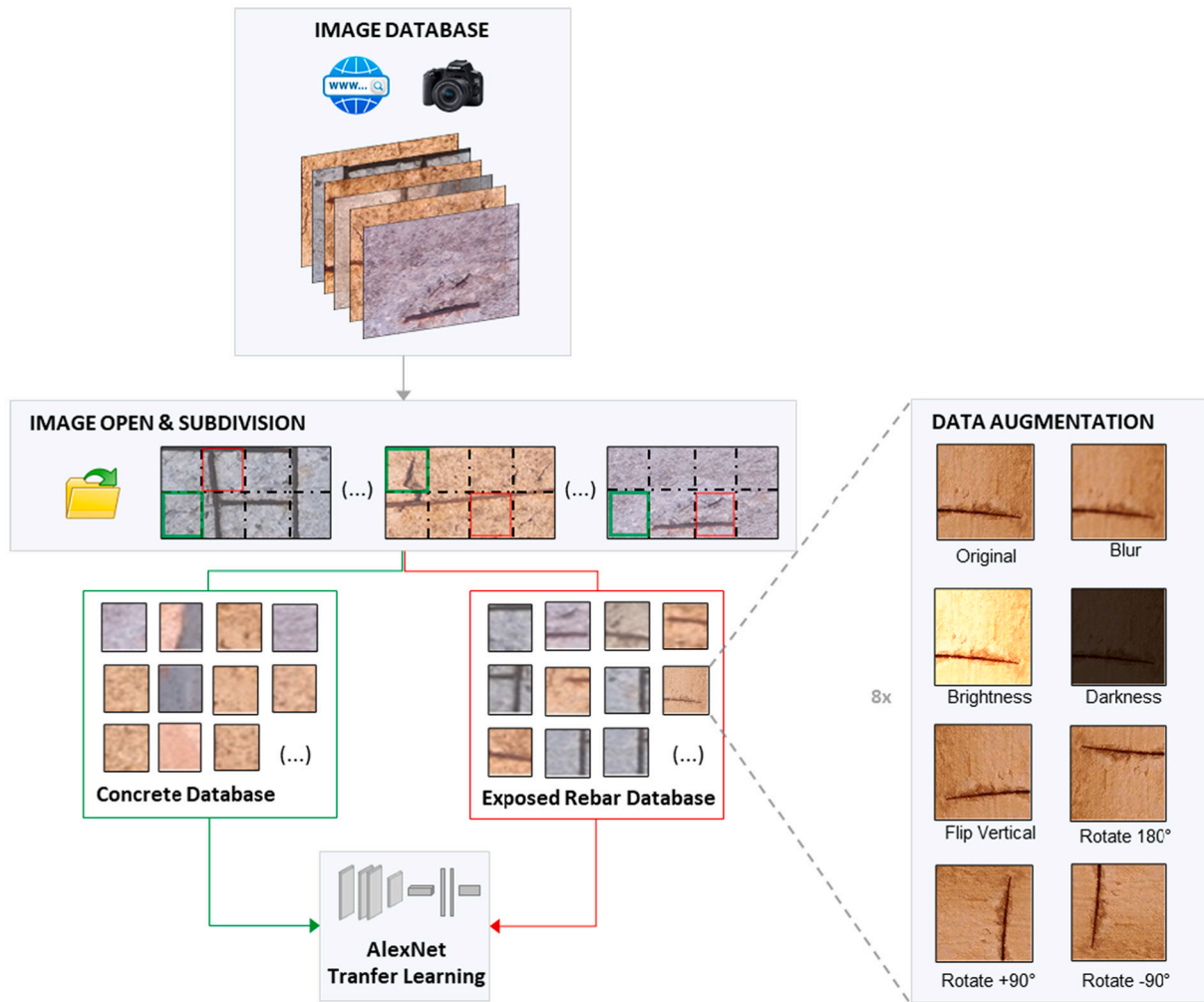
Fig. 2. Methodology description.

Fig. 3. Flowchart of the Image database constitution.

Image classification involves assigning class labels to an image, whereas object identification involves highlighting one or more objects in an image. A typical CNN can only determine the class of objects found in an image but not their location [30]. To overcome this challenge, Girshick *et al* [31] proposed a method that combine regions with CNNs, called R-CNNs. In R-CNNs the CNN is forced to focus on a single region of the larger image at each given time, carried on by a selective search technique for object detection, followed by a resizing algorithm forcing regions to be of equal size before they are fed to a R-CNN for object identification.

However, research using CNNs normally require large amounts of carefully chosen data to train a network. When the amount of the training data is insufficient underfitting will likely occur, conducting to erroneous or sub-standard results [32]. Transfer learning techniques attempt to transfer knowledge from previous consolidated and trained CNNs to aid targeting new tasks when the latter has fewer high-quality training data thus avoiding possible underfitting problems [33].

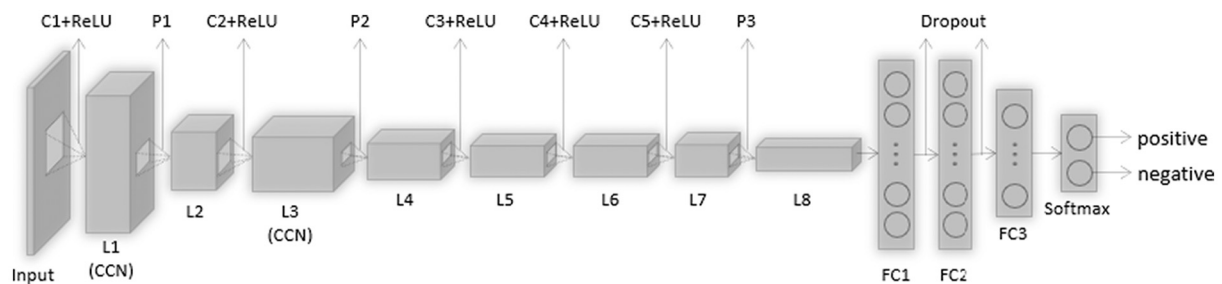AlexNet is a well-trained CNN, based on more than a million images

Fig. 4. Network architecture. L#: Operation layers; C#: Convolution operations; P#: Max pooling operations; CCN: Cross-channel normalization; FC#: Fully connected layers.

from the ImageNet database [34]. The network has a depth of eight layers, the first five being convolutional layers, and the last three fully connected layers. AlexNet can classify images into 1000 object categories, including keyboards, mouses, pencils, and many animals. The network has an image input size of 227 × 227 pixels.

Dorafshan et al [35] show the superiority of the artificial intelligence approach using AlexNet network with transfer learning mode over digital image processing techniques for concrete crack detection. On their work the dataset was composed of 3420 sub-images of concrete cracks from a handheld camera. The digital image processing technique accurately detected about 79% of concrete cracks and was useful in detecting cracks coarser than 0.1 mm. In comparison, the AlexNet network, in transfer learning mode, accurately detected 86% of crack images and could detect cracks coarser than 0.04 mm. AlexNet network was chosen to develop this work due to its easy implementation after the releasing of so many deep learning libraries. Rubio et al [36] based on a database of images from a handheld camera of bridges in Niigata Prefecture, Japan, developed a Fully Convolutional Network (FCN), a dense CNN without fully connected layers using VGG-16 network as the backbone, for detecting delamination and exposed steel rebars, showing a mean accuracy of 89.7% and 78.4%, respectively. Shin et al [11] proposed a model that can automatically classify different shapes and types of damage that occur on concrete surfaces in real conditions. To prepare the training dataset, Shin et al classified the data from a handheld camera into the following categories: crack (530), delamination (563), rebar exposure (268), leakage (208), as well as non-damage (412) images, through the Concrete Multi-Damage recognition neural network (CMDnet) designed on a VGG-16 network. Concrete surface multi-damage recognition using CMDnet achieved a 95.7% minimum probability of correct prediction, with a 98.9% accuracy.

Advances in unmanned aerial systems (UAS) have produced low-cost and high-mobility UAVs, expanding their use on real-world applications including the civil engineering field [37–40]. UAVs allow surface visual inspections in large-scale civil engineering structures, permitting better results on the identification of anomalies [23]. In recent years, the application of Digital image processing techniques and CNNs in data images collected by UAVs, assisted on the implementation of efficient visual inspections, which have begun to attract researchers for more robust non-contact damage detection [41–45].

Ribeiro et al [23] describe a methodology for remote inspection of reinforced concrete structures using UAVs with the integration of results into BIM models based on advanced digital image processing techniques and the application of heuristic method for feature extraction. Alternatively,

Rajadurai et al [43] focused on deep convolutional neural networks to detect and classify cracks through images taken from UAV by using transfer learning from AlexNet network. The trained model presented a prediction accuracy of 99.9%. Chaiyasarn et al [44] propose an automated crack detection system based on CNNs, along with UAV technology, applied in damage inspection. The 3D models created from the acquired and processed images can be used to archive structure properties, as well as to create an inspection report in a 3D model form with

associated image locations to provide a better view for inspectors. Kumar et al [45] use a real-time multi-UAS approach to detect concrete's cracks and spalls, offering reliable performance with an accuracy of 94.24% and an image processing in 0.033 s.

The present article proposes an innovative contribution to some aspects: the creation of a CNN exclusively dedicated to the identification of exposed steel rebars based on a large number of images (20*k*) with variable lighting exposures, simultaneously integrating the use of UAVs in the methodology of exposed steel rebar detection. Two large structures under real environmental conditions, namely an industrial building and a telecommunications tower, are presented as case studies. Additionally, the authors have created orthoimage mosaics with georeferenced identification of the surface anomalies, a feature not included on any document mentioned in the consulted bibliography.

2. Methodology for automatic detection of exposed steel rebars

2.1. General description

The proposed methodology for automatic detection of exposed steel rebars in RC structures is based on a video system installed on an UAV, also known as drone, supported by advanced image processing techniques based on Artificial Intelligence (AI) (Fig. 1). The images collected by the UAV video system are georeferenced based on the drone built-in GPS, which may eventually be aided by a Real Time Kinematic (RTK) position correction system, should the need arise. The AI image processing techniques are based on Convolutional Neural Networks (CNNs) that can detect exposed steel rebars on each processed image. On a second phase, a georeferenced orthophoto mosaic were built merging the several individual images. The UAV video system, together with the CNN image postprocessing, proved to be robust enough to provide successful results under the external environmental conditions existing during the image acquisition campaigns, particularly under low/medium winds and variable lightning exposure including shadows.

2.2. Equipment

Table 1 illustrates the specifications of two UAVs from DJI, which were used as part of the system components in the proposed methodology.

The DJI Matrice 600 Pro is a hexacopter with a maximum recommended take-off weight of 15.5 kg, a maximum horizontal speed of 65 km/h, a maximum service ceiling of 2500 m above sea level and an estimated flight autonomy of approximately 32 min (per battery set). The camera installed in the drone has a gimbal mounted DJI Zenmuse X5, with a 4/3" CMOS sensor and a resolution of 16 Megapixels. The gimbal ensures image stabilization along the three axes. The visualization of the images and the remote control of the drone were carried out by means of a DJI Matrice 600 Series remote controller and a high-resolution tactile DJI Crystal Sky monitor, with internal image storage capacity.

The DJI Mavic Mini is a quadcopter with a maximum recommended

Table 2

CNN layer description.

| Layers / Description	Schematic representation

INPUT: At the beginning of the process, the input image is divided into pixels. For a black and white image, each pixel is assigned a value between 0 and 255, corresponding to the black and white colours respectively, where in between these values there is a grey scale. A colour image is represented by a three-dimensional matrix, so that it is possible to store the combination of the three colours of the RGB scale.

[Input Image]   **[Input Matrix]**
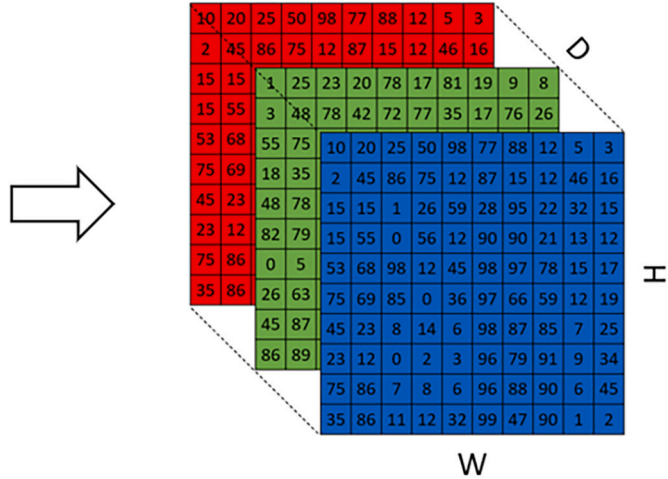
CONVOLUTION: The purpose of convolution is to extract features from the input image. It is a mathematical operation that has as input data, an image matrix and a filter matrix. The filter matrix runs through the entire image matrix, resulting in a smaller dimension matrix that will be easier to process. Convolution of an image with different filters can identify characteristics such as colours, gradient, orientation and borders.

[Input Matrix]   **[Filter Matrix]**   **[Output Matrix]**
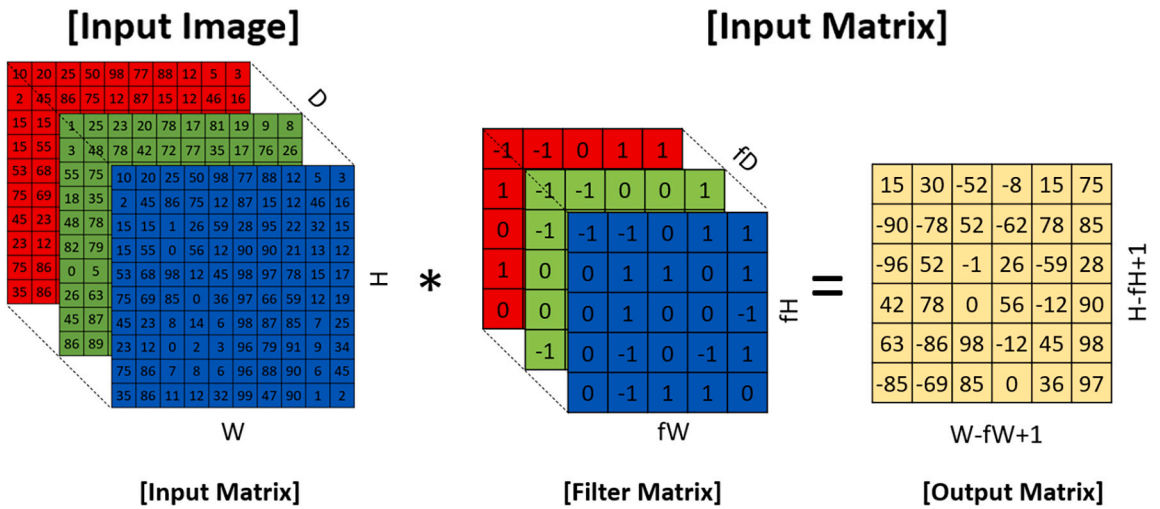
RELU: ReLU stands for "rectified linear units". In this layer the function f (x) = max (0,x) is applied. The purpose of this layer is to introduce non-linearity to the network. Since most complex problems are nonlinear, it is important to use this activation function.

[Convolved Matrix]   **[RELU Matrix]**
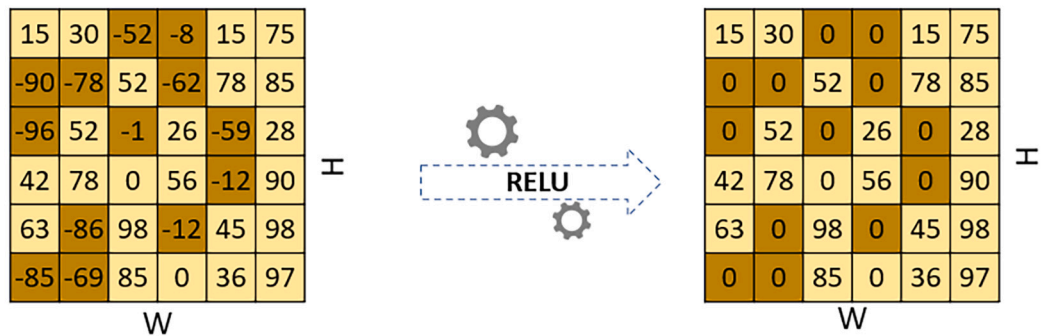
POOLING: After extracting features through the convolution layer, the resulting feature map is passed to the Pooling layer immediately after

Table 2 (*continued*)

Layers / Description	Schematic representation
applying a rectified linear activation function. This layer reduces the dimension of the input, extracting the most dominant information, so it is possible to reduce the necessary computational load, maintaining the CNN training process efficiency. The most used types are Max Pooling and Average Pooling.	
FULLY CONECTED LAYER: This layer is used for the classification of the characteristics extracted in the previous layers. Resultant maps are converted to vectors and their characteristics are combined to create a classifier.	

take-off weight of 249 g, a maximum horizontal speed of 14 km/h, a maximum service ceiling of 3000 m above sea level and an estimated flight autonomy of approximately 30 min (per battery set). The drone has a native gimbal mounted camera, with internal stabilizers that cannot be disabled and a 12 Megapixels resolution 1/2.3" CMOS sensor.

For the studies conducted, the precision obtained by the native GNSS systems on both UAVs used was deemed sufficient, since all image capture campaigns were conducted on open-wide areas with a robust GNSS signal. Should the need arise, other vehicles with RTK assisted geolocation might be used. Presently, this technology is not available on any of the drones used (DJI Mavic Mini and DJI Matrice 600 Pro).

3. Region convolutional neural network (R-CNN) for exposed steel rebar detection

3.1. Process description

The process of creating an efficient neural network has two essential steps: the constitution of a database of images and the training and validation of the network. The quality of the database is probably the most important factor and should cover, whenever possible, the most varied scenarios that can be found in a real environment. Hence, a well-graded diversified database has much more chances of success on the detection and identification of whatever features it is designed for when applied to real situations.

Another important decision to consider is the choice between a pre-trained network or the creation of a network from scratch. The disadvantage of creating one's own network is the need for a very large database to achieve a minimally satisfactory efficiency rate, which is why pre-trained networks are usually chosen. There are several pre-trained networks of different architectures that can be implemented, and, as a rule, the most efficient ones require more processing capacity, leading to longer training times. The training parameters also have some influence on the final efficiency, as these parameters must include the percentage of the database that the user intends to reserve for training and validation, the learning rate, the maximum number of epochs and the batch size. An entire dataset that is passed forward and backward through the neural network is considered an epoch. Since one epoch is normally too big to be fed to the computer on one single instance it is normally divided into several smaller batches, called mini-batches.
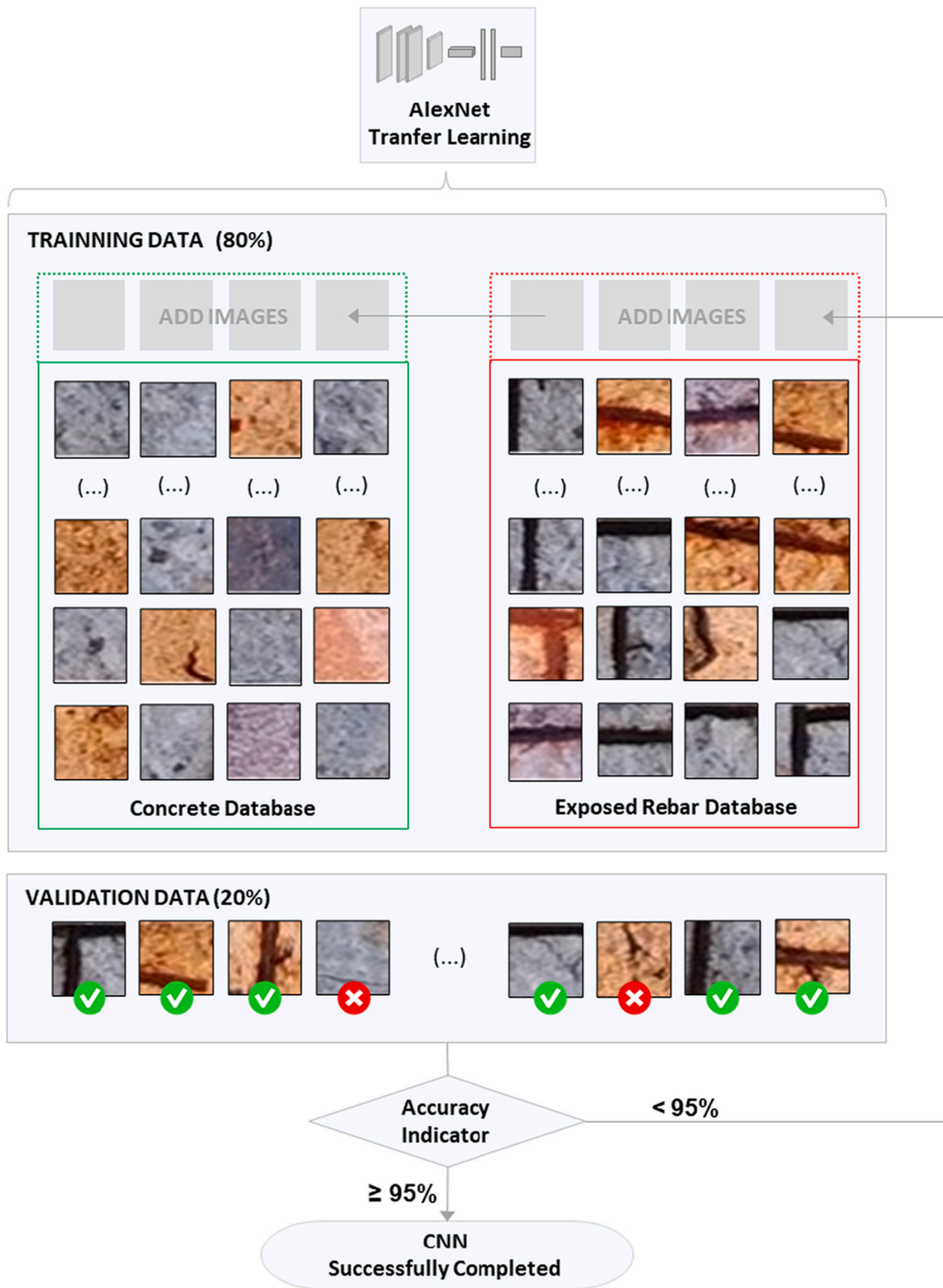
Fig. 5. CNN training and validation flowchart.

With the network chosen well-trained, an automatic damage detection can be performed as illustrated in Fig. 2. However, the methodology proposed in this work consists on the application of a neural network for the detection of exposed steel rebars in a real environment, using the AlexNet pre-trained network and a database developed by the authors. High resolution close captured images of the entire structure under inspection were used and processed through a sliding window by the trained CNN. The sliding window parsed each image 227 × 227 pixels at a time, darkening the image in regions where no exposed rebars were found and keeping the remaining areas unaltered. The images were subsequently stitched together to form a georeferenced map highlighting the anomalies found.
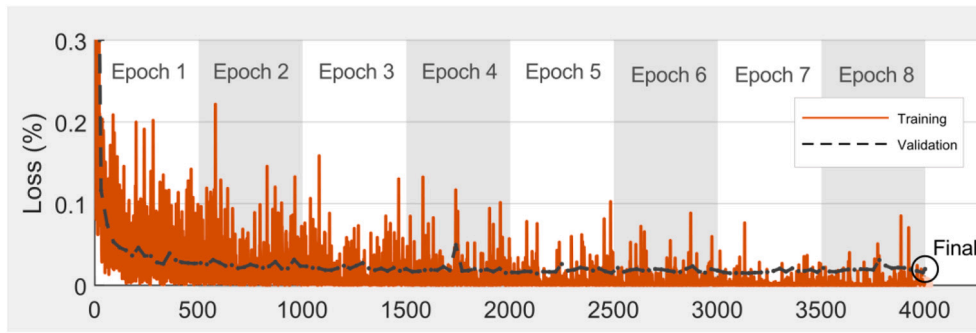
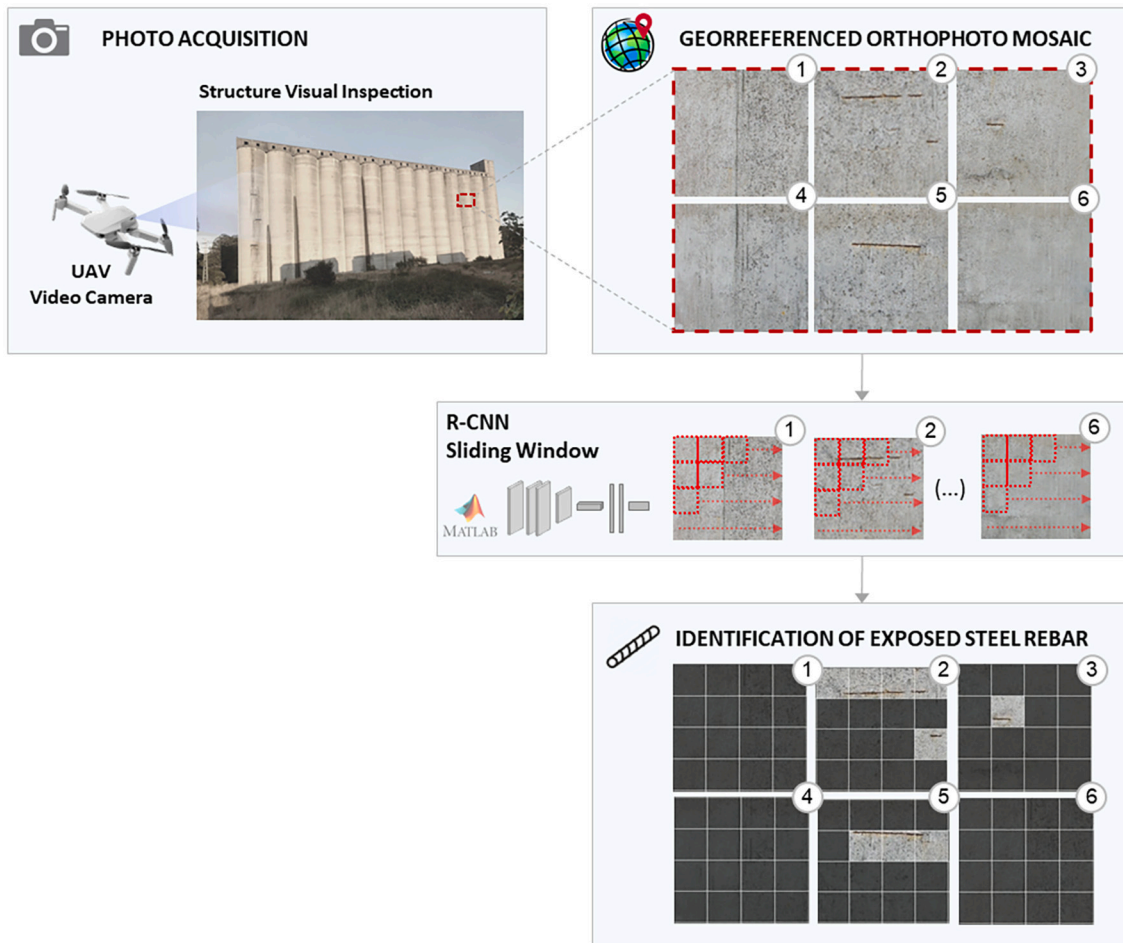**Fig. 6.** Training loss and validating loss over training epochs.

Fig. 7. Trained CNN application on reinforced concrete structures

In terms of performance evaluation of CNNs, four metrics are normally used (Accuracy, Precision, Recall and F1 Score). Accuracy (Eq. (1)) is a statistical measure which is defined as the quotient of correct predictions made divided by the sum of all predictions. Precision (Eq. (2)) is the ratio of the correctly identified positive cases to all the predicted positive cases. Recall (Eq. (3)), also known as sensitivity, is the ratio of the correctly identified positive cases to all the actual positive cases. At last, F1 score (Eq. (4)) conveys the balance between the precision and the recall.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \times 100 \qquad (1)$$

$$Precision = \frac{TP}{TP + FP} \times 100 \qquad (2)$$

$$Recall = \frac{TP}{TP + FN} \times 100 \qquad (3)$$

- FP (False Positives) indicates the number of no exposed steel rebars that are falsely detected as exposed rebars;
- FN (False Negative) indicates the number of exposed steel rebars that the algorithm did not recognize as exposed rebars.

3.2. Image database

The development of the exposed steel rebar automatic detection tool, started with the acquisition of images of both pathology ridden and intact concrete. It involved capturing images in real context and obtaining online images. The images obtained in real context were captured using a smartphone camera with a resolution of 4000 × 3000 pixels, a DJI Mavic Mini drone camera with a resolution of 4000 × 2250 pixels, a DJI Zenmuse X5 camera with a resolution of 4608 × 3452 and also a Canon EOS 250d professional camera with a resolution of 4000 × 3000 pixels. The images were captured from several structures in the district of Porto, Portugal, and taken at several angles, distances, and different times of the day in order to cover a large diversity of image capture situations.

Images were then subdivided into 227 × 227 pixel sub images as this is the pre-defined image input for the pre-trained network used in this work. All images which were not considered valid for classification after division were discarded.

Finally, it was also necessary to use Data Augmentation techniques in order to obtain enough images for a robust classifier. Training images can be increased in number by geometric transformations, blur and changes in the colour scale. As for geometric transformations, rotations, translations and reflections can be applied to the original image in order to take into account the variation of the direction and the angle at which an image can be obtained. The application of blur simulates image capture conditions of insufficient light or motion, while changes in the colour scale, are intended to take into consideration variations in white balance or brightness and colour of the pictures.

The applied transformations allowed a seven-fold increase of the original database size, obtaining a sufficient number of valid images for training the CNN. At the end of the process, 40 k images were obtained, 20 k of exposed reinforcement and 20 k of intact concrete. Fig. 3 illustrates the process of the database constitution.

3.3. Training and validation

The Deep Learning Toolbox from Matlab©, in which some of the most popular convolutional neural network algorithms are implemented, was used to perform the CNN training. CNNs use 2D convolutional layers which makes its architecture ideal for processing 2D data such as images.

Additionally, CNNs eliminate the need to perform a manual feature extraction, doing it automatically through the provided images. This automatic extraction makes the Deep Learning model highly reliable for object classification and require less pre-processing when compared to other image classification algorithms. The example of a CNN architecture applied to the recognition of characteristics in images can be found on Fig. 4.

In the Matlab© Deep Learning Toolbox there are 20 algorithms of convolutional neural networks embedded, including the most popular like AlexNet, GoogleNet, Resnet and VGG-19 networks. These networks differ from each other in the design of their architectures, varying the number of layers, parameters, size of the input data and disk occupation resulting in networks that behave differently and consequently give different results, with different training times and problem solution precision.

CNNs use the predictions from the layers to produce a final output that presents a vector of probability scores to represent the likelihood that a specific feature belongs to a certain class.

Besides the input, there are four types of layers for AlexNet and most of the CNNs used today: the convolutional layer, the pooling layer, the
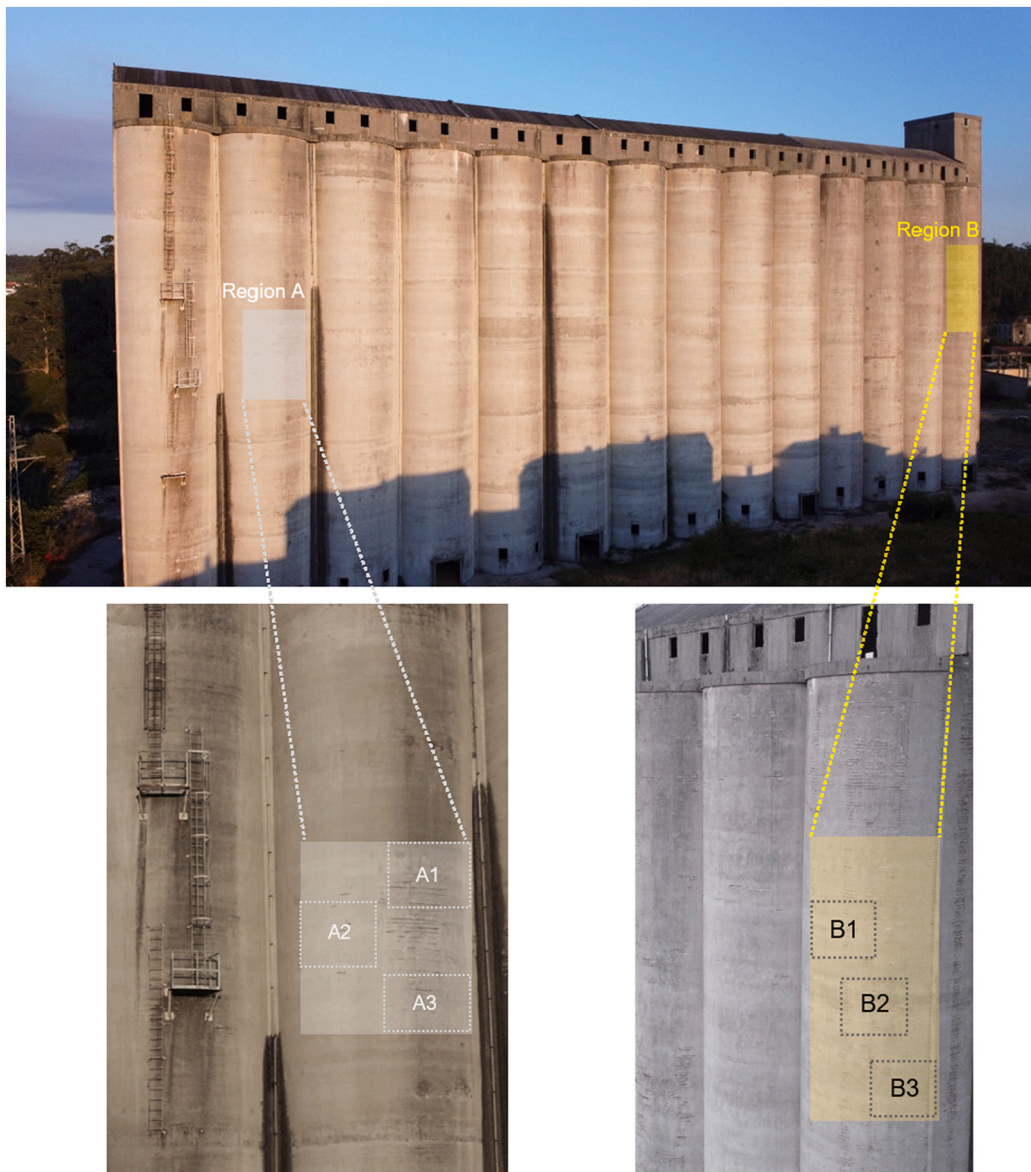

Fig. 8. Industrial building, view from the bridge over the Ave River.

Fig. 9. Aerial view of the studied structure.

$$F1\ Score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \times 100 \qquad (4)$$

where:

- TP (True Positives) indicates, in this case, the number of exposed steel rebars successfully detected by the algorithm;
- TN (True Negative) indicates the number of no exposed steel rebars successfully detected by the algorithm;

Fig. 10. Identification of the inspection regions.

ReLU correction layer and the fully connected layer. Table 2 describes the transformations that occur in the layers.

It was decided to use the AlexNet pre-trained network for the training of the CNN. In the absence of previous collected data for this particular problem, the decision for choosing this network was based on the analysis of studies in which similar problems are solved. Özgenel & Sorguç [46] did a comparison of the performance of seven highly acknowledged pretrained convolutional neural networks on crack detection in buildings. Tests were performed with Alexnet, GoogleNet, ResNet50, ResNet101, ResNet152, VGG16 and VGG19. The results reveal that all networks had scored over 90% in accuracy, with AlexNet scoring in the middle tier of these results. However, on what concerns the training time, AlexNet revealed to be over nine times faster than the

second fastest option. Similar results are reported in [45].

As the CNN AlexNet is configured for one thousand categories, it was necessary to make small changes to its layers in order to classify only two categories. For each category, the images were subdivided, with 80% used for training the CNN and 20% reserved for validation. Fig. 5 shows, schematically, how the process of training a robust classifier works.

The training was performed on a medium range computer with the following characteristics: Processor: i5–7500, Graphics Card: GTX 1050 2GB, Motherboard: ASUS H110M-K, RAM: HyperX Fury DDR4 8GB (2400 MHz), Hard Drives: Kingston SSD 120GB and TOSHIBA HDD 1 TB and OS: Windows 10 64bit. The network training took about 8 h and resulted in an accuracy of 99.1%. As shown in Fig. 6, solid orange light curves and dashed dark black lines represent the training loss and
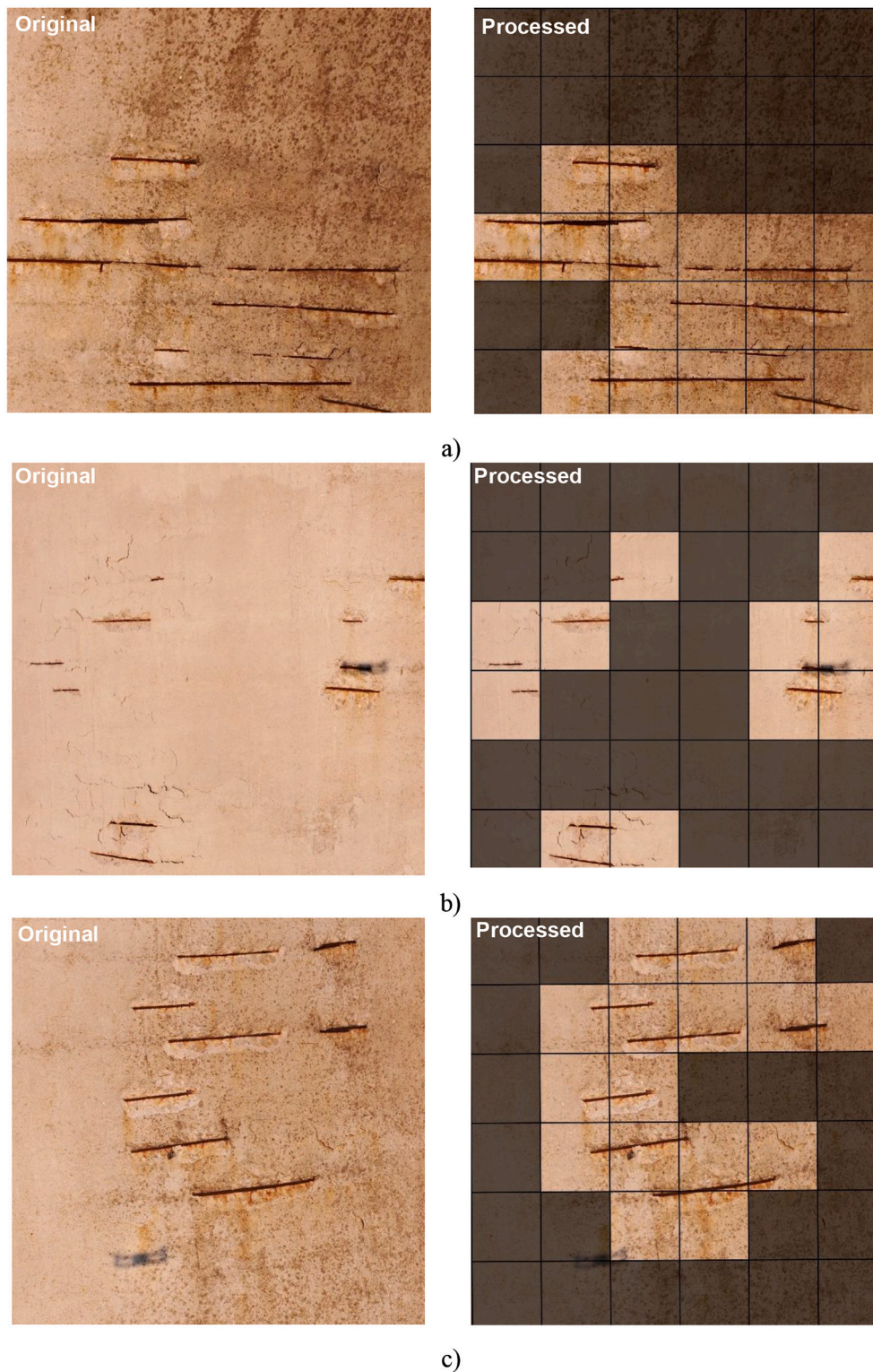
a)

b)

c)

Fig. 11. Image processing in Region A. a) subregion A1; b) subregion A2; c) subregion A3.

validating loss, respectively during the interactions into the 8 epochs. In this case, the validating loss performed close to the training loss to ensure that the parameters of the CNN were not overfit during the training progress. The metrics discussed above (see Section 3.1), namely, the accuracy, precision, recall, and F1 score achieved results of 97.6%, 98.9%, 97.4%, and 98.1%, respectively.

3.4. Application and georeferenced images with marked steel rebars

Having a CNN well-trained with satisfactory results, it became necessary to develop an application to evaluate test images in order to prove the real efficiency of the CNN. An algorithm was developed in Matlab© capable of processing input images, regardless of the original resolution, by processing consecutive fractions of each image. As explained before, a sliding window traverses the image at strides of 227
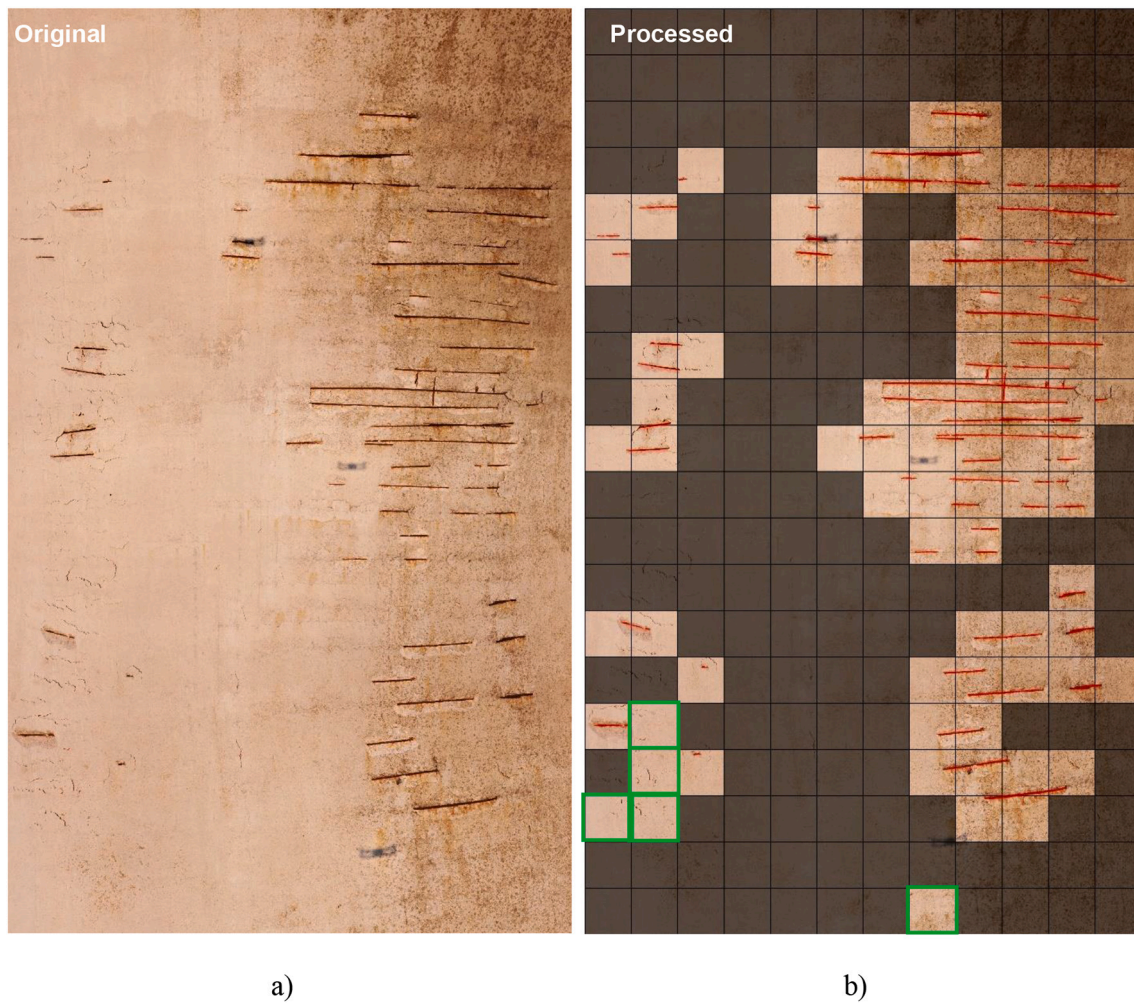
Fig. 12. Inspection of region A: original stitched image *vs* stitched image with marked steel rebars.

pixels, both horizontally and vertically, darkening the image in the fractions where the classifier does not find anomalies, thus highlighting the areas with anomalies in the processed image.

The various images captured in the studied structures were processed and stitched to form a georeferenced map with the anomalies found, using the software Image Compositor Editor©. The process of creating the georeferenced map is illustrated in Fig. 7.

4. Case studies

4.1. Silo structure

4.1.1. Description

The first case study is an industrial building, located in Vila do Conde, Portugal, next to the bank of the Ave River and about two kilometres from the Atlantic Ocean. It is a large building, about 20 m wide, 100 m long and 35 m high, built entirely in reinforced concrete (Fig. 8). The building, now abandoned, belonged to the company "Prazol", that was once dedicated to grinding, treating, and storing of seeds and

cooking oils. The company went into decline at the end of the 20th century and was closed before the end of the century. The industrial complex in which the studied building is integrated, is now in an advanced state of degradation.

4.1.2. Application

Image capturing was carried out in July 2020, in the late afternoon. The conditions for using the UAV were close to ideal, since it was a sunny day with calm winds, and with plenty of light. For security reasons, entrance to the industrial complex was forbidden, due to the advanced state of degradation of the structure and facilities. This limitation determined that the capture zone should be limited to the West side of the structure in order to safely operate the drone. Fig. 9 shows an aerial view of the structure with the position of the drone pilots and the locations chosen for image capturing.

On the silo, images were captured in two different locations, both in the West side of the structure. Fig. 10 shows the inspection regions where the images were captured. From left to right, on the second silo cell the grey rectangle marks Region A and on the last silo cell the yellow
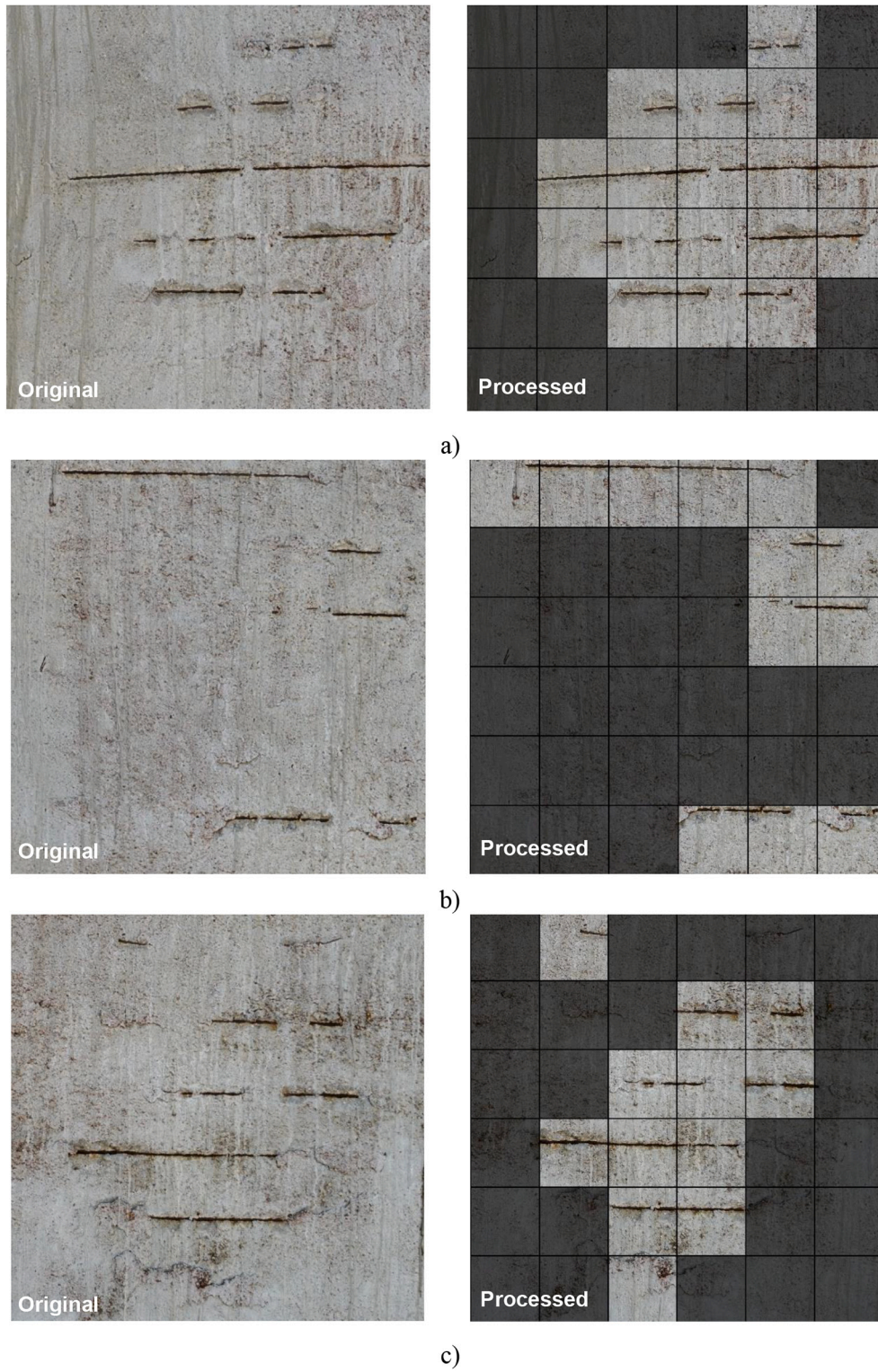
Fig. 13. Image processing in Region B. a) subregion B1; b) subregion B2; c) subregion B3.
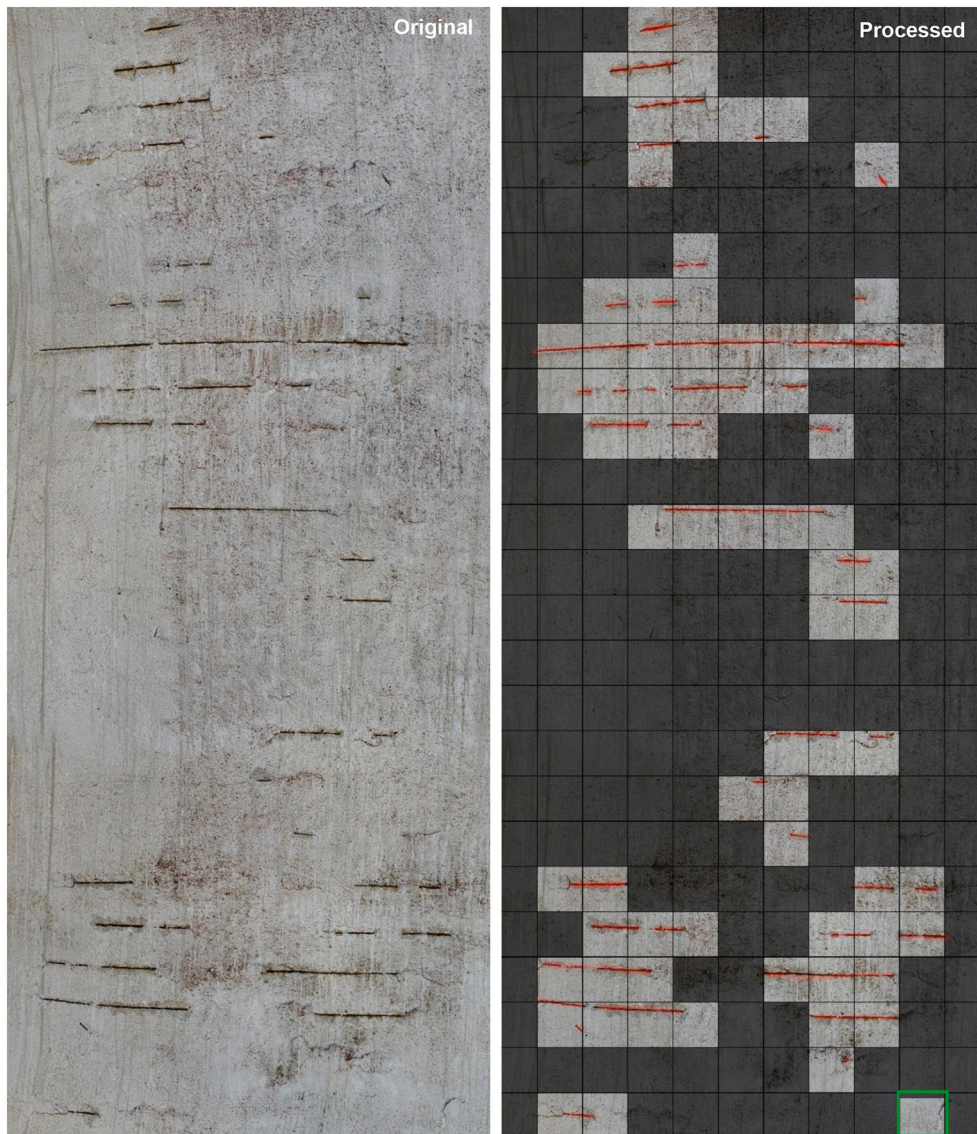
Fig. 14. Inspection of region B: original stitched image *vs* stitched image with marked steel rebars.

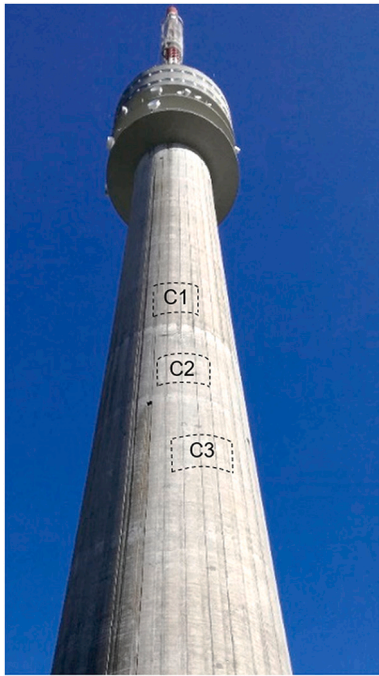**Fig. 15.** Aerial view of the Telecommunications tower.

Fig. 16. Telecommunication tower inspection sub-regions.

rectangle marks Region B.

Fig. 11 presents three examples of the application of the exposed rebar detection tool applied to images on region A.

Fig. 12 shows the aggregated image before and after processing, in which false positives found are highlighted in green. Exposed rebars have been highlighted in red for increased visibility.

In total, 4 false positives were found in the stitched image, with the accuracy of the network in this image being 98.33%, which is higher than the validation efficiency. The false positives found appear in areas of advanced delamination, but still without exposed rebars, leading the network to erroneously classify these positions.

Fig. 13 shows again three examples of the application of the same tool applied to Region B.

Fig. 14 shows the aggregated images before and after processing, in which the false positives found are again highlighted in green and the exposed rebars highlighted in red.

In this stitched image, only one false positive was found, with the network accuracy in this image being 99.60%, which is also higher than the validation efficiency. Apart from a small crack in the right side of the inspection region, there are no obvious reasons why the network failed in this position. The most likely cause appears to be that the network has mixed up the small dark spot of the crack with the beginning of an exposed rebar.

4.2. Telecommunications tower

4.2.1. Description

The Monte da Virgem telecommunications tower is a transmission tower built by Altice/Portugal Telecom in 1995 and located in Vila Nova de Gaia, in the north of Portugal. The structure of the tower consists of a 126 m high RC shaft and a 51 m high metallic mast, for a total height of 177 m. It is the highest structure of its kind in Portugal (Fig. 15).

The RC shaft has the shape of a hyperboloid of revolution, with a circular hollow section and a diameter varying between 14.3 m, at the base, and 7.7 m at the top. The shaft includes five technical floors, materialized by prestressed concrete cantilever slabs. The technical floors are located between the heights of 94.9 m and 112 m, measured to the base of the shaft, with two of these closed with external walls, while the three remaining floors are open and protected with balcony railings. The metallic mast on top consists of a spatial lattice composed of three sections along its height.

4.2.2. Application

Similar to the silo case study, the photographs were taken at the communications tower and processed afterwards. However, the number of photographs was not sufficient to map a continuous region of interest, mostly due to flight conditions hampered by strong winds at the moment of the photograph capture. Three photographs are presented for the application of the tool on the tower. Fig. 16 frames the images in the structure while Fig. 17 shows the obtained results.

5. Conclusions

This article describes the implementation of a technique for contactless inspection of reinforced concrete structures using unmanned aerial vehicles, with the purpose of automatically detect exposed reinforcement rebars and create orthoimage mosaics with georeferenced identification of this pathology on concrete wall images.

The proposed technique relies on a two-stage strategy where, on the first stage, images are acquired using drones in order to create a set of images that will later be merged on an orthomosaic representation of a building façade. On a second stage, images are analysed using artificial intelligence algorithms based on deep learning, which, in this particular implementation, rely on Region-Convolutional Neural Networks (R-CNNs) based on the AlexNet CNN. This approach has the advantage of being able to automatically identify the existence or absence of exposed steel rebars and simultaneously obtain a representation of their location superimposed on the orthomosaic of the façade.

The algorithm was implemented using the Deep Learning Toolbox from Matlab© and required the acquisition and processing of a large number of images (20 K images) in order to build a well-graded, diversified image database covering various scenarios found in a real environment. This image database served both as the training and validation core for the applied neural network.

The metrics applied to evaluate the performance of the CNN based on the validation subset composed of about 20% of the number of collected images, included measurements on the accuracy, precision, sensitivity (normally named Recall) and balance between precision and sensitivity (normally named F1 score). These metrics achieved consistent results of 97.6%, 98.9%, 97.4%, and 98.14%, respectively, for each of the previously mentioned criteria. The training process resulted on a 99.1% accuracy.

For the two case studies presented, a silo and a telecommunications tower, and for the regions analysed in each case, the accuracy obtained was consistent, if not better, than the results of the evaluation metrics applied.

For future developments, the authors are considering the advantages of upgrading the proposed methodology replacing the sliding windows method with a selective search or image pyramids for object detection. Other improvements considered include the instance segmentation of multiple concrete damage based on Mask R-CNN and U-Net techniques. Additionally, some specific field applications will be carried out to evaluate the potential for metric characterization of anomalies.
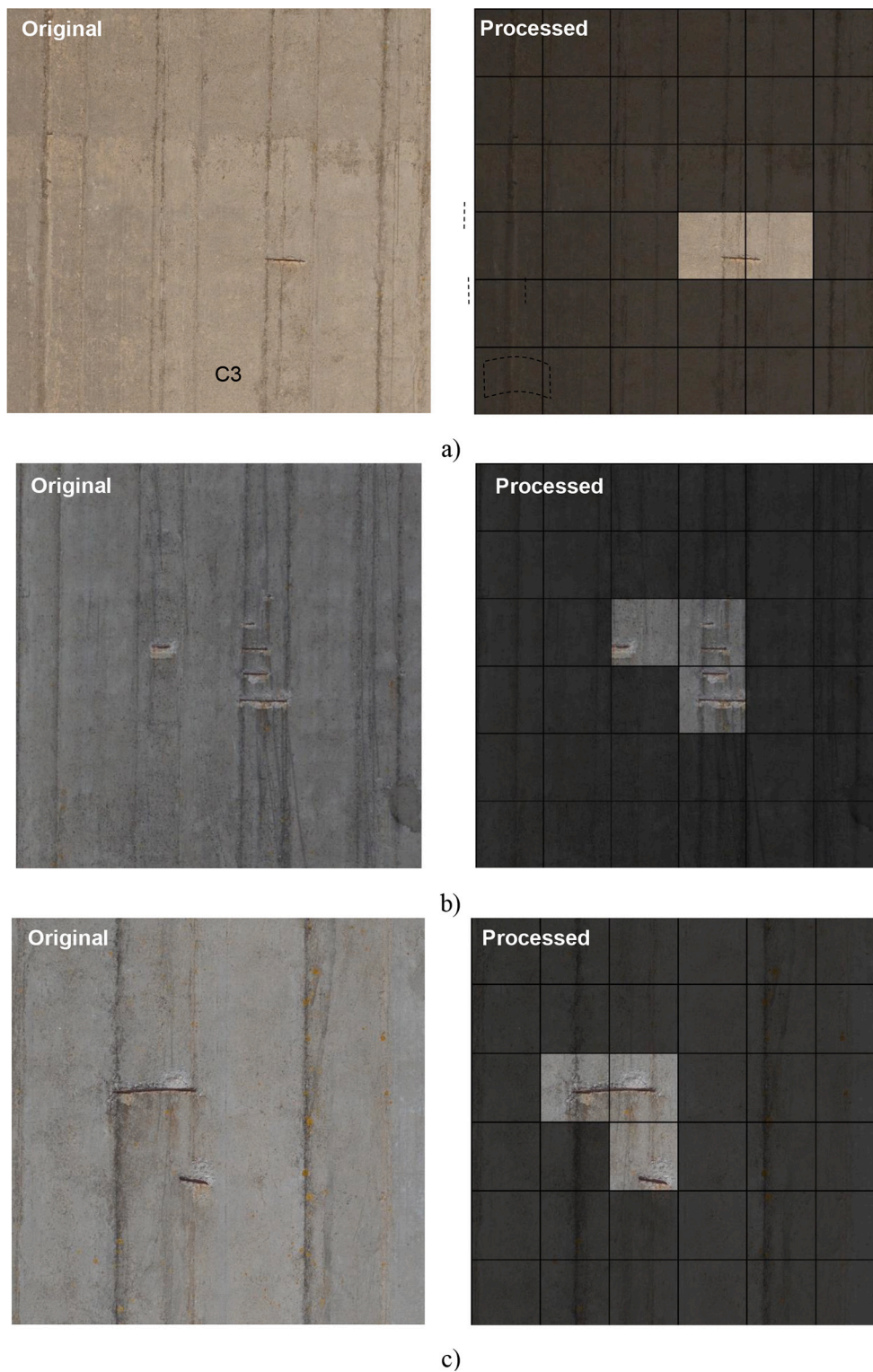
Fig. 17. Inspection subregions: a) subregion C1, b) subregion C2, c) subregion C3.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

MCTES.

References

[1] A. Momayez, M. Ehsani, A. Ramezanianpour, H. Rajaie, Comparison of methods for evaluating bond strength between concrete substrate and repair materials, Cem. Concr. Res. 35 (4) (2005) 748–757, https://doi.org/10.1016/j.cemconres.2004.05.027.

[2] Y.-J. Cha, W. Choi, Vision-based concrete crack detection using a convolutional neural network, in: Dynamics of Civil Structures Vol. 2, Springer, 2017, pp. 71–73, https://doi.org/10.1007/978-3-319-54777-0_9.

[3] C.V. Dung, L.D. Anh, Autonomous concrete crack detection using deep fully convolutional neural network, Autom. Constr. 99 (2019) 52–58, https://doi.org/10.1016/j.autcon.2018.11.028.

[4] P.J. Chun, S. Izumi, T. Yamane, Automatic detection method of cracks from concrete surface imagery using two-step light gradient boosting machine, Comp. Aided Civil Infrastruct. Eng. 36 (1) (2021) 61–72, https://doi.org/10.1111/mice.12564.

[5] N. Wang, X. Zhao, P. Zhao, Y. Zhang, Z. Zou, J. Ou, Automatic damage detection of historic masonry buildings based on mobile deep learning, Autom. Constr. 103 (2019) 53–66, https://doi.org/10.1016/j.autcon.2019.03.003.

[6] D.J. Atha, M.R. Jahanshahi, Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection, Struct. Health Monit. 17 (5) (2018) 1110–1128, https://doi.org/10.1177/1475921717737051.

[7] M. Khayatazad, L. De Pue, W. De Waele, Detection of corrosion on steel structures using automated image processing, Dev. Built Environ. 3 (2020), 100022, https://doi.org/10.1016/j.dibe.2020.100022.

[8] C. Cheng, Z. Shang, Z. Shen, Automatic delamination segmentation for bridge deck based on encoder-decoder deep learning through UAV-based thermography, NDT & E Int. 116 (2020), https://doi.org/10.1016/j.ndteint.2020.102341.

[9] A.A. Ijjeh, S. Ullah, P. Kudela, Full wavefield processing by using FCN for delamination detection, Mech. Syst. Signal Process. 153 (2021), 107537, https://doi.org/10.1016/j.ymssp.2020.107537.

[10] B. Kim, S. Cho, Automated multiple concrete damage detection using instance segmentation deep learning model, Appl. Sci. 10 (22) (2020), https://doi.org/10.3390/app10228008.

[11] H.K. Shin, Y.H. Ahn, S.H. Lee, H.Y. Kim, Automatic concrete damage recognition using multi-level attention convolutional neural network, Materials 13 (23) (2020), https://doi.org/10.3390/ma13235549.

[12] S.K. Verma, S.S. Bhadauria, S. Akhtar, Monitoring corrosion of steel bars in reinforced concrete structures, Sci. World J. 2014 (2014), 957904, https://doi.org/10.1155/2014/957904.

[13] A.N. Kalashnikov, R.E. Challis, Errors and uncertainties in the measurement of ultrasonic wave attenuation and phase velocity, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 52 (10) (2005), https://doi.org/10.1109/TUFFC.2005.1561630.

[14] V. Marcantonio, D. Monarca, A. Colantoni, M.J.M.S. Cecchini, S. Processing, Ultrasonic waves for materials evaluation in fatigue, thermal and corrosion damage: A review 120, 2019, pp. 32–42, https://doi.org/10.1016/j.ymssp.2018.10.012.

[15] F. Ghodoosi, A. Bagchi, T. Zayed, M.R. Hosseini, Method for developing and updating deterioration models for concrete bridge decks using GPR data, Autom. Constr. 91 (2018) 133–141, https://doi.org/10.1016/j.autcon.2018.03.014.

[16] H. Ahmed, H.M. La, K. Tran, Rebar detection and localization for bridge deck inspection and evaluation using deep residual networks, Autom. Constr. 120 (2020), https://doi.org/10.1016/j.autcon.2020.103393.

[17] S. Doshvarpassand, C. Wu, X. Wang, An overview of corrosion defect characterization using active infrared thermography, Infrared Phys. Technol. 96 (2019) 366–389, https://doi.org/10.1016/j.infrared.2018.12.006.

[18] G. Duffó, N. Gaillard, M. Mariscotti, M. Ruffolo, Application of gamma-ray radiography and gravimetric measurements after accelerated corrosion tests of steel embedded in mortar, Cem. Concr. Res. 74 (2015) 1–9, https://doi.org/10.1016/j.cemconres.2015.03.017.

[19] S. Kim, J. Surek, J. Baker-Jarvis, Electromagnetic metrology on concrete and corrosion, Res. Nat. Inst. Stand. Technol. 116 (3) (2011) 655–669, https://doi.org/10.6028/jres.116.011.

[20] T. Nishikawa, J. Yoshida, T. Sugiyama, Y. Fujino, Concrete crack detection by multiple sequential image filtering, Comp. Aided Civil Infrastruct. Eng. 27 (1) (2012) 29–47, https://doi.org/10.1111/j.1467-8667.2011.00716.x.

[21] T.H. Dinh, Q.P. Ha, H.M. La, Computer vision-based method for concrete crack detection, in: 2016 14th international conference on control, automation, robotics and vision (ICARCV), IEEE, 2016, pp. 1–6, https://doi.org/10.1109/ICARCV.2016.7838682.

[22] E. Protopapadakis, A. Voulodimos, A. Doulamis, N. Doulamis, T. Stathaki, Automatic crack detection for tunnel inspection using deep learning and heuristic image post-processing, Appl. Intell. 49 (7) (2019) 2793–2806, https://doi.org/10.1007/s10489-018-01396-y.

[23] D. Ribeiro, R. Santos, A. Shibasaki, P. Montenegro, H. Carvalho, R. Calçada, Remote inspection of RC structures using unmanned aerial vehicles and heuristic image processing, Eng. Fail. Anal. 117 (2020), https://doi.org/10.1016/j.engfailanal.2020.104813.

[24] P. Prasanna, K.J. Dana, N. Gucunski, B.B. Basily, H.M. La, R.S. Lim, H. Parvardeh, Automated crack detection on concrete bridges, IEEE Trans. Autom. Sci. Eng. 13 (2) (2014) 591–599, https://doi.org/10.1109/TASE.2014.2354314.

[25] Y.-J. Cha, W. Choi, O. Büyüköztürk, Deep learning-based crack damage detection using convolutional neural networks, Comp. Aided Civil Infrastruct. Eng. 32 (5) (2017) 361–378, https://doi.org/10.1111/mice.12263.

[26] B. Kim, S. Cho, Automated vision-based detection of cracks on concrete surfaces using a deep learning technique, Sensors (Basel) 18 (10) (2018), https://doi.org/10.3390/s18103452.

[27] L. Guo, R. Li, B. Jiang, X. Shen, Automatic crack distress classification from concrete surface images using a novel deep-width network architecture, Neurocomputing 397 (2020) 383–392, https://doi.org/10.1016/j.neucom.2019.08.107.

[28] C. Su, W. Wang, Z.-Z. Wang, Concrete cracks detection using convolutional neural network based on transfer learning, Math. Probl. Eng. 2020 (2020) 1–10, https://doi.org/10.1155/2020/7240129.

[29] H.-N. Li, L. Ren, Z.-G. Jia, T.-H. Yi, D.-S. Li, State-of-the-art in structural health monitoring of large and complex civil infrastructures, Civil Struct. Health Monit. 6 (1) (2016) 3–16, https://doi.org/10.1007/s13349-015-0108-9.

[30] C. Modarres, N. Astorga, E.L. Droguett, V. Meruane, Convolutional neural networks for automated damage recognition and damage type identification, Struct. Control. Health Monit. 25 (10) (2018), e2230, https://doi.org/10.1002/stc.2230.

[31] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (2014) 580–587, https://doi.org/10.1109/CVPR.2014.81.

[32] H. Kim, H. Kim, Y.W. Hong, H. Byun, Detecting construction equipment using a region-based fully convolutional network and transfer learning, Comp. Civ. Eng. 32 (2) (2018), https://doi.org/10.1061/(asce)cp.1943-5487.0000731.

[33] Q. Yang, W. Shi, J. Chen, W. Lin, Deep convolution neural network-based transfer learning method for civil infrastructure crack detection, Autom. Constr. 116 (2020), https://doi.org/10.1016/j.autcon.2020.103199.

[34] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Proces. Syst. 25 (2012) 1097–1105, https://doi.org/10.1145/3065386.

[35] S. Dorafshan, R.J. Thomas, M. Maguire, Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete, Constr. Build. Mater. 186 (2018) 1031–1045, https://doi.org/10.1016/j.conbuildmat.2018.08.011.

[36] J.J. Rubio, T. Kashiwa, T. Laiteerapong, W. Deng, K. Nagai, S. Escalera, K. Nakayama, Y. Matsuo, H. Prendinger, Multi-class structural damage segmentation using fully convolutional networks, Comput. Ind. 112 (2019), https://doi.org/10.1016/j.compind.2019.08.002.

[37] H. Kim, J. Lee, E. Ahn, S. Cho, M. Shin, S.H. Sim, Concrete crack identification using a UAV incorporating hybrid image processing, Sensors 17 (9) (2017), https://doi.org/10.3390/s17092052.

[38] X. Zhong, X. Peng, S. Yan, M. Shen, Y. Zhai, Assessment of the feasibility of detecting concrete cracks in images acquired by unmanned aerial vehicles, Autom. Constr. 89 (2018) 49–57, https://doi.org/10.1016/j.autcon.2018.01.005.

[39] S. Sreenath, H. Malik, N. Husnu, K. Kalaichelavan, Assessment and use of unmanned aerial vehicle for civil structural health monitoring, Proc. Comp. Sci. 170 (2020) 656–663, https://doi.org/10.1016/j.procs.2020.03.174.

[40] H.-F. Wang, L. Zhai, H. Huang, L.-M. Guan, K.-N. Mu, G.-P. Wang, Measurement for cracks at the bottom of bridges based on tethered creeping unmanned aerial vehicle, Autom. Constr. 119 (2020), https://doi.org/10.1016/j.autcon.2020.103330.

[41] K. Chaiyasarn, Crack detection in historical structures based on convolutional neural network, Int. J. GEOMATE 15 (51) (2018), https://doi.org/10.21660/2018.51.35376.

[42] I.H. Kim, H. Jeon, S.C. Baek, W.H. Hong, H.J. Jung, Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle, Sensors 18 (6) (2018), https://doi.org/10.3390/s18061881.

[43] R.S. Rajadurai, S.T. Kang, Automated vision-based crack detection on concrete surfaces using deep learning, Appl. Sci. 11 (11) (2021) 5229, https://doi.org/10.3390/s18103452.

[44] K. Chaiyasarn, M. Sharma, L. Ali, W. Khan, N. Poovarodom, Crack detection in historical structures based on convolutional neural network, Int. J. GEOMATE 15 (51) (2018), https://doi.org/10.21660/2018.51.35376.

[45] P. Kumar, S. Batchu, S.R. Kota, Real-time concrete damage detection using deep learning for high rise structures, IEEE Access 9 (2021) 112312–112331, https://doi.org/10.1109/ACCESS.2021.3102647.

[46] Ç.F. Özgenel, A.G. Sorguç, 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, https://doi.org/10.22260/ISARC2018/0094.