




Using digital image processing and Unmanned Aerial Vehicle (UAV) for identifying ceramic cladding detachment in building facades

Uso de processamento digital de imagens e Veículo Aéreo Não-Tripulado (VANT) na identificação de descolamento cerâmico em fachadas de edifícios

Alcineide Dutra Pessoa de Sousa 
Gean Carlos Lopes de Sousa 
Luiz Maurício Furtado Maués 

Abstract

Facade maintenance actions are driven by results obtained in the inspection phase. Some methodological proposals aimed at optimizing the inspection process have been discussed, notably digital image processing (DIP) techniques associated with unmanned aerial vehicle (UAV) imagery. Using UAV speeds up the access to the inspected area, and DIP techniques help to automate the identification of pathological manifestations. This article aims to apply DIP techniques to detect areas where the ceramic cladding on building facades is detaching. The methodology referred to herein starts with the creation of a database (images) captured by cell phone and UAV. The object detection algorithm YOLO (You Only Look Once) was applied to the database images. The results indicated these techniques are very promising, with a 94% precision level in the tests performed. The precision index obtained indicates that the model is applicable in practice and discussions about its limitations help improve the proposed methodology.

Keywords: Façade inspection. Digital image processing. UAV. YOLO. Ceramic detachment.

Resumo

As ações que compõem o processo de manutenção em fachadas são orientadas pelos resultados obtidos na fase de inspeção. Algumas propostas metodológicas que visam a otimização do processo de inspeção têm sido discutidas, entre essas destacam-se as técnicas de Processamento Digital de Imagens (PDI) capturadas por Veículo Aéreo Não Tripulado (VANT). O uso de VANT para capturar imagens em fachadas agiliza o acesso à área inspecionada, e as técnicas de PDI ajudam a automatizar o processo de identificação de manifestações patológicas. Esse artigo tem como objetivo aplicar técnicas de PDI para detectar regiões com descolamento cerâmico em fachadas de edifícios. A metodologia utilizada inicia com a formação de uma base de dados (imagens) capturadas por celular e por VANT. O algoritmo de detecção de objetos YOLO (You Only Look Once) foi aplicado nas imagens que compõem a base de dados. Os resultados obtidos indicaram que as técnicas utilizadas são bastante promissoras, com precisão de 94% nos testes realizados. O índice de precisão obtido indica que o modelo é aplicável na prática e as discussões sobre as limitações ajudam no melhoramento da metodologia proposta.

Palavras-chave: Inspeção da fachada. Processamento digital de imagens. VANT. YOLO. Descolamento cerâmico.

¹Alcineide Dutra Pessoa de Sousa

¹Universidade Federal do Pará
Belém - PA - Brasil

²Gean Carlos Lopes de Sousa

²Universidade Federal do Maranhão
Imperatriz - MA - Brasil

³Luiz Maurício Furtado Maués

³Universidade Federal do Pará
Belém - PA - Brasil

Recebido em 11/05/21

Aceito em 04/10/21

Introduction

In order to preserve or restore the functional capacity of a building and its parts, several actions can be taken. The set of such actions is called maintenance and is aimed at providing for the necessities and safety of the user and should be driven by an inspection process to assess the state of the building and its parts (ABNT, 2012).

In the specific case of building facades, the inspection process becomes even more meticulous because some variables such as height, access difficulties and exposure conditions must be considered. Depending on these variables, the task of the facade inspection can become laborious, expensive, and even dangerous (RUIZ *et al.*, 2021). Several methodologies have been proposed for performing the facade inspection (KAVUMA; OCK; JANG, 2019). Ruiz *et al.* (2021), however, point out that visual inspection using an unmanned aerial vehicle (UAV) is an effective and safe method, and also reduces the time and cost of inspection.

The process that combines visual inspection and the use of a UAV can be improved through computer vision techniques (RUIZ *et al.*, 2021). Several studies have sought to identify the different types of pathological manifestations on building facades (GALANTUCCI; FATIGUSO, 2019; PIRES; DE BRITO; AMARO, 2015).

A systematic bibliographic survey was conducted in this research using the Start software, whereby the main scientific databases were investigated; furthermore, the systematic review by Pan and Zhang (2021) was used, where the authors listed 30 articles on computer vision applied to civil engineering issues. The papers found both in the bibliographic survey and in the systematic review by Pan and Zhang (2021) demonstrate that using artificial intelligence applied to engineering problems, especially those related to civil engineering, spells for a promising research field.

In this context of applying image processing techniques in civil engineering issues, particularly concerning pathology identification, Valero *et al.* (2019) used machine learning to detect defects in masonry walls. Perez, Tah and Mosavi (2019) used Convolutional Neural Networks (CNN) to identify problems such as mold, stains, and paint deterioration in images obtained through various sources. Another task that has been the subject of studies is the detection of cracks in concrete structures (SONG *et al.*, 2019; LI; ZHAO, 2019; ISLAM; KIM, 2019).

Artificial intelligence techniques, especially computer vision techniques, are commonly applied to images captured in a variety of ways. However, one methodology that has stood out in image capture is the use of unmanned aerial vehicle (UAV) (KERLE *et al.*, 2019). According to Kerle *et al.* (2019), the use of images obtained through UAV allows for the reconstruction of highly detailed and accurate scenes, which, in turn, allows for creating 3D models (JEONG *et al.*, 2020; CHEN *et al.*, 2021). Several types of pathologies have been identified through these models, such as those existing in facades with ceramic tiles (BALLESTEROS; LORDSLEEM JUNIOR, 2021).

Considering that using UAV images and artificial intelligence techniques have provided important results when applied to various types of engineering issues, the herein research introduces a methodology for automated detection of ceramic tile detachment on building facades. This methodology combines artificial intelligence and the use of UAV to collaborate with the evaluation process and, consequently, with the maintenance of building facades.

Identification of pathological manifestations on building facades

According to Madureira *et al.* (2017), building facades are a complex system to design, build and maintain. This system consists of walls, openings and different types of claddings. Over time, pathological issues often occur. These problems, however, can be attenuated and/or solved through detailed inspections, that is, by effectively identifying pathological manifestations (SILVESTRE; DE BRITO, 2011).

Silvestre and De Brito (2011) proposed a methodology for identifying anomalies in ceramic cladding applied to facades. This methodology was based on standardized inspections and was applied to 85 ceramic cladding systems along with a quantitative analysis of the types of pathological manifestations.

In order to facilitate the evaluation of facade degradation, especially painted ones, Pires, De Brito and Amaro (2015) systematized methods of inspection, diagnosis and rehabilitation through the compilation and structuring of data contained in regulations and in scientific publications, in addition to techniques based on visual inspection or data compilation, as in the cases by Silvestre and De Brito (2011) and Pires, De Brito and Amaro (2015), the identification of pathologies on facades using image processing has gained space.

The methodologies using digital image processing and UAV have proved to be effective, safe and more economical (BALLESTEROS; LORDSLEEM JUNIOR, 2021; RUIZ *et al.*, 2021). In order to identify the knowledge gap in this area, a Systematic Literature Review (SLR) was conducted between March and August/2020. Conducting a SLR was based on the search for articles through strings formed from the following keywords: Unmanned Aerial Vehicle (UAV), Convolutional Neural Network (CNN), Building, Image processing, Construction pathologies, ceramic detachment, Efflorescence, facades, Spalling, Hole, Crack. This search was carried out on the basis of articles of Engineering Village, Scopus, IEE, Science Direct and Web of Science, resulting in a set of 2540 articles found.

Out of the 2540 articles found in the databases, 460 duplicates were removed or had import errors, thus 2080 articles were left. Out of these, 2062 were excluded due to the established exclusion criteria (articles unrelated to the theme, conference articles and articles older than 5 years from publication), thus leaving 15 works directly related to the identification of pathologies using Digital Image Processing and UAV. Table 1 summarizes the articles on identifying pathological manifestations found in SLR.

According to the information in Box 1, the most recent articles (2019 and 2020) are those that used machine learning technology or UAV to identify pathologies. There is yet a need to develop research on the triple relationship “use of UAV – use of CNN – Ceramic detachment”. It is in this context that the herein research proposes the use and analysis of a methodology that combines image processing and UAV to detect ceramic cladding detachment in building facades.

Convolutional neural networks

Artificial neural networks (ANN) are computer algorithms capable of learning patterns. These algorithms/models are inspired in the biological behavior of the human brain. ANN models are used in various applications, such as digital image processing (DA COSTA; DE LIMA; BARBOSA, 2021). The most commonly used neural networks in digital image processing (DIP) are the so-called Convolutional Neural Networks (CNN) (KATTENBORN *et al.*, 2021).

When the DIP is intended for object identification, the YOLO (You Only Look Once) network has stood out (BARREIROS *et al.*, 2021). Performing DIP using CNN has culminated in a diversity of applications, such as detecting human face (ALI-GOMBE *et al.*, 2021), objects on rural roads (BARBA-GUAMÁN *et al.*, 2021), underwater objects (AYOB *et al.*, 2021); construction worker safety hats (ZHANG *et al.*, 2021) and even for monitoring mask use in the times of pandemic (LOEY *et al.*, 2021). The layers of a YOLO-like CNN are composed of four types of operations, namely: convolution, max pooling, rule, and batch normalization.

Table 1 - Works found on the identification of pathological manifestations in buildings: CNN - Convolutional Neural Networks, UAV - unmanned aerial vehicle

Authors	CNN	UAV	Pathology
Valero <i>et al.</i> (2019)	X		Defects in masonry
Song <i>et al.</i> (2019)	X		Cracks
Li and Zhao (2019)	X		Cracks
Yu <i>et al.</i> (2019)	X		Structural damage
Islam and Kim (2019)	X		Cracks
Perez, Tah and Mosavi (2019)	X		Mould and stains
Morillas <i>et al.</i> (2015)			White efflorescence
Pires, De Brito and Amaro (2015)			Paintings on facades
Galantucci and Fatiguso (2019)			Cracks
Bauer, Castro and Silva (2015)			Detachment
Diaz, Cornadó and Albareda (2020)			Brick damage
Bauer, Milhomem and Aidar (2018)			Cracks in facades
Kerle <i>et al.</i> (2019)		X	Structural damage
Pereira and Pereira (2015)		X	Cracks
Sreenath <i>et al.</i> (2020)		X	Structural damage

Convolution operation is responsible for filtering the images. This process aims to extract inherent characteristics of the data presented to the network. To this end, spatial filters are used in the convolution operation. A convolution in \mathbf{R}^2 is defined as (Equation 1):

$$(f * g)(x, y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i, j)g(x - i, y - j) \tag{Eq. 1}$$

Where:

f is the original image;

g is the filter; and

indexes i and j correspond to pixel positions.

The goal of the max pooling operation is to reduce the size of the activation map through subsampling. The most widely adopted subsampling methodology is the maximum value approach (Figure 1).

The operation known as ReLU aims to replace all negative values with zero. An element-by-element applied activation function is used for this task. This function is defined as follows (ALHEEJAWI *et al.*, 2020) (Equation 2):

$$f = \max(0, y) \tag{Eq. 2}$$

Where y is the input element for ReLU and f is the output.

The Batch Normalization operation performs normalization of the input x_i elements. This normalization is performed by considering the mean μ and variance σ^2 over the spatial dimensions for each channel independently. Equation xx indicates how the batch normalization operation is performed (IOFFE; SZEGEDY, 2015) (Equation 3).

$$x_{norm} = \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}} \tag{Eq. 3}$$

Where x_{norm} is the normalized sample and ϵ is a constant that is directly linked to numerical stability, especially when variance is close to zero.

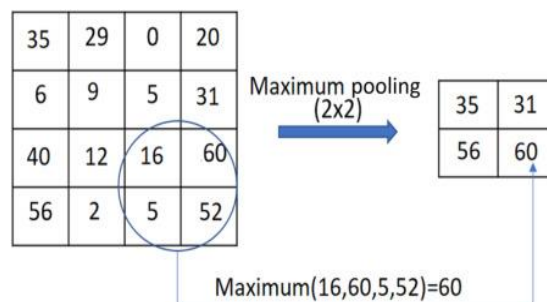
The way the operations described above are organized in the structure of a YOLO-type CNN can be seen in Figure 2.

In the last layer of the YOLO (the Anchor Layer), the objects to be identified are detected. In this layer, the bounding boxes that accommodate the objects are defined, such as the region where ceramic cladding was detached. It is worth noting that in addition to the bounding box, the YOLO algorithm presents the probability that the investigated object is contained in that box.

Materials and method

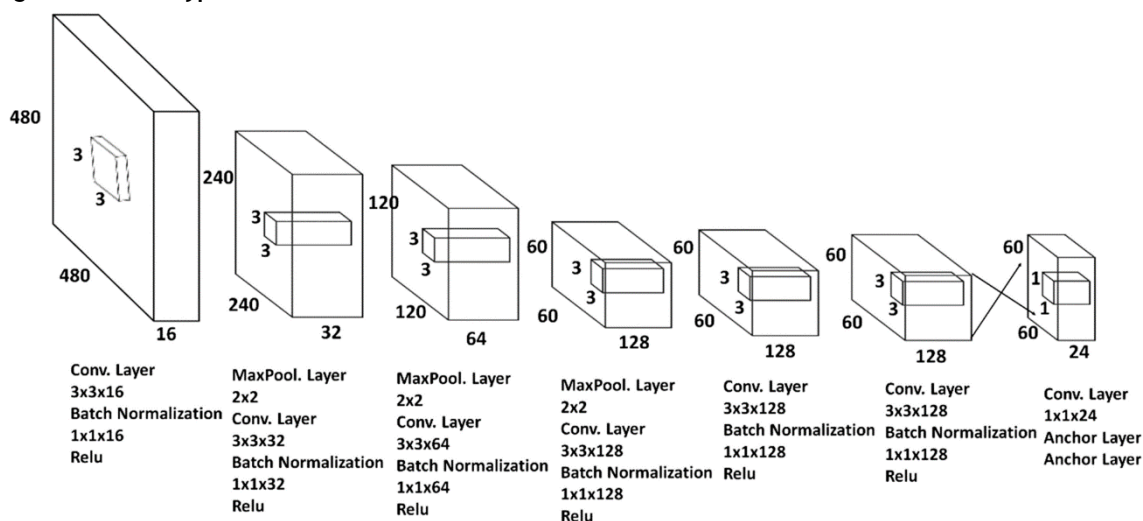
The methodology applied herein consisted of two phases (Figure 3). The first phase involves the creation of a bank of images to be processed in a later instance. The second phase consisted of two stages: pre-processing and application of the YOLO algorithm for detecting the detachment of ceramic cladding in images/videos of facades. According to Barreiros *et al.* (2021), the YOLO network is an easy-to-implement algorithm and has a higher performance (precision) than other networks. Besides, this algorithm can perform object detection in real time. The distinctive good performance and detection speed of YOLO were the main reasons that led us to use this algorithm in the methodology proposed herein.

Figure 1 - Example of maximum value subsampling



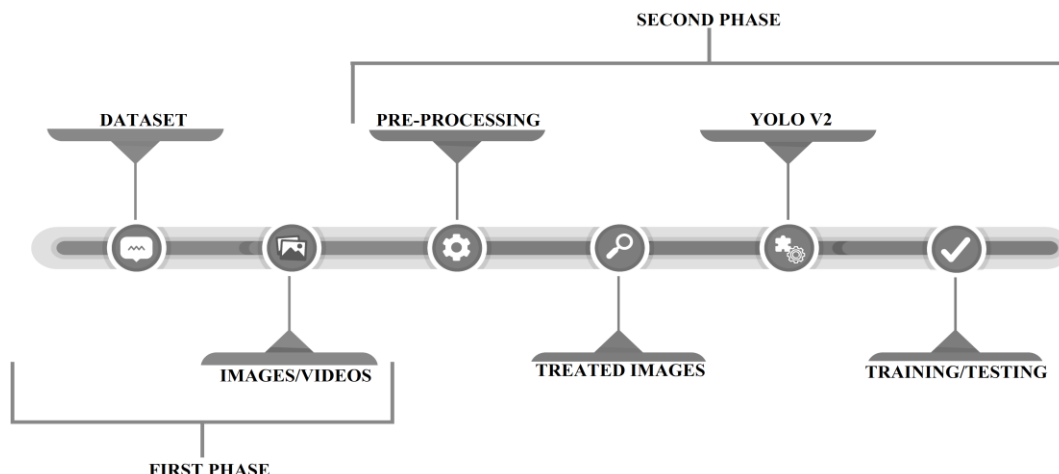
Source: Singh and Majumder (2020).

Figure 2 - YOLO-type CNN architecture



Source: Barreiros *et al.* (2021).

Figure 3 - Phases of the work methodology



Typically, for most artificial intelligence algorithms, the YOLO v2 network needs to be used in two phases, training and testing (the latter is shown in Figure 3). In the training phase, images are presented so that the network can learn the patterns referring to the problem under study. In the test phase, new images are presented to the network and the number of hits and errors are counted in order to measure the performance thereof.

Dataset construction

The statement of Ruiz *et al.* (2021) that there is a lack of public dataset for this type of problem, besides being true, demonstrates how important it is to obtain this information and that it is a challenge for researchers to obtain images to make up the dataset. Thus, the herein research started by creating a dataset in order to train and test the model being used. It is worth mentioning that good results are directly proportional to the amount of images used for training the neural network, that is, the composition of the dataset can hinder good results.

To try to overcome the possible limitations arising from the number of images in the dataset, two actions were taken. The first was to use a pre-processing methodology to increase the number of images. The other was to use a network with pre-trained weights. According to Yosinski *et al.* (2014), using these weights reduces the need to retrain all the parameters of the neural network.

The dataset was then constructed and it covered three image/video subsets:

- (a) captured cell phone images of the facades of historical, commercial and public buildings in the business center of the cities of Belém (in Pará State), São Luís and Imperatriz (both in Maranhão State);
- (b) drone images of the facade of two residential buildings in Belém; and
- (c) a drone video of the facade of a third residential building in Belém.

At the end of the image/video collection task, a dataset of 687 images and 1 video was obtained. Eighty percent (80%) of the images were randomly separated for training and 20% of them were for testing. The facade video obtained with the UAV in the construction of the dataset was also used for testing the detector.

Image pre-processing

Before applying the YOLO v2 object identification algorithm, the images must be labeled. The training and testing images were labeled manually using the Image Labeler application from MathWorks company. Amongst other functions, this application has often been employed for this type of task (MATHWORKS, 2020a). Figure 4 shows an example of an image with and without labeling.

After labeling all images, the YOLO v2 network was applied to identify areas of cladding detachment and to measure the correctness of the network. The application of this algorithm and the parameters being used are described in the following topic.

Object detection using the YOLO v2 network

Deep learning consists in a data processing technique that has been the subject of recent studies. This technique allows for training robust object detectors, such as the YOLO network. Currently, the YOLO network is implemented in several versions with improvements especially related to speed issues. The second version of this network is known as YOLO v2 (PLASTIRAS; KYRKOU; THEOCHARIDES, 2019; REDMON; FARHADI, 2017).

The YOLO v2 model uses a deep-learning convolutional network to extract characteristics from the input images. These characteristics are then decoded and generate bounding boxes (Figure 5).

Figure 4 - Example of a photo pre-processed by the Image Labeler application

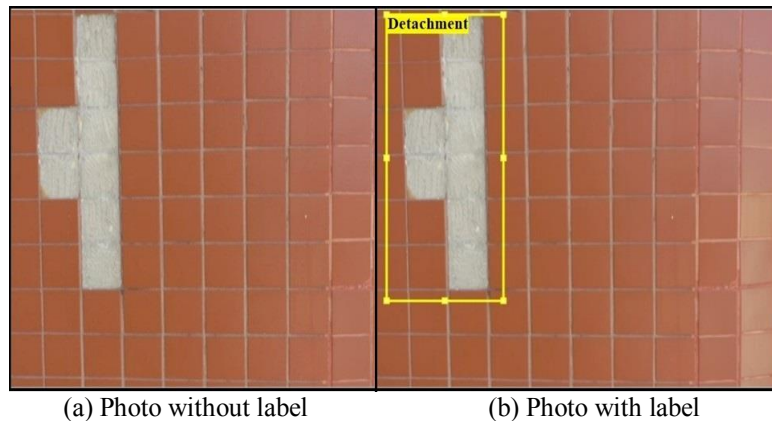
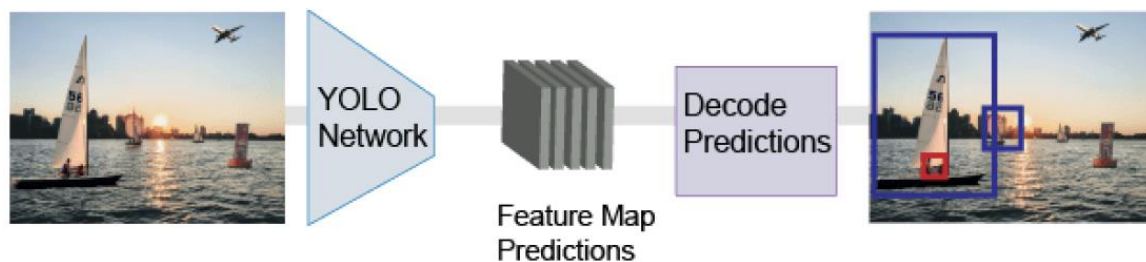


Figure 5 - Detecting objects with YOLO



Source: MathWorks (2020b).

The YOLO v2 convolutional network is capable of detecting objects with higher performance than conventional detection methods (BARREIROS *et al.*, 2021). The YOLO v2 algorithm divides the image into several grids that detect an object within the grid. These grids are called anchor boxes.

For each anchor box, YOLO v2 provides the class label assigned to each anchor box and a confidence index, among other things. This index is the probability for the object to fall into the bounded box. Then, YOLO v2 tries to eliminate, as much as possible, the bounding boxes that do not correspond to the class of the object (BARREIROS *et al.*, 2021; REDMON; FARHADI, 2017). In short, the scores encode the probability of each class and how good the predicted box fits into the object (Figure 6).

The architecture of the Yolo network is, in general, made up of a set of layers, whereby each layer performs a specific function. The typically used layers are: input, Addition, Batch Normalization, Convolutional; Max Pooling, ReLU, YOLO v2 Transform, and output. The input layer corresponds to the network input; the Addition layers add subsequent layers. YOLO's convolutional layers decrease the sample by a factor of 32 (BARREIROS *et al.*, 2021).

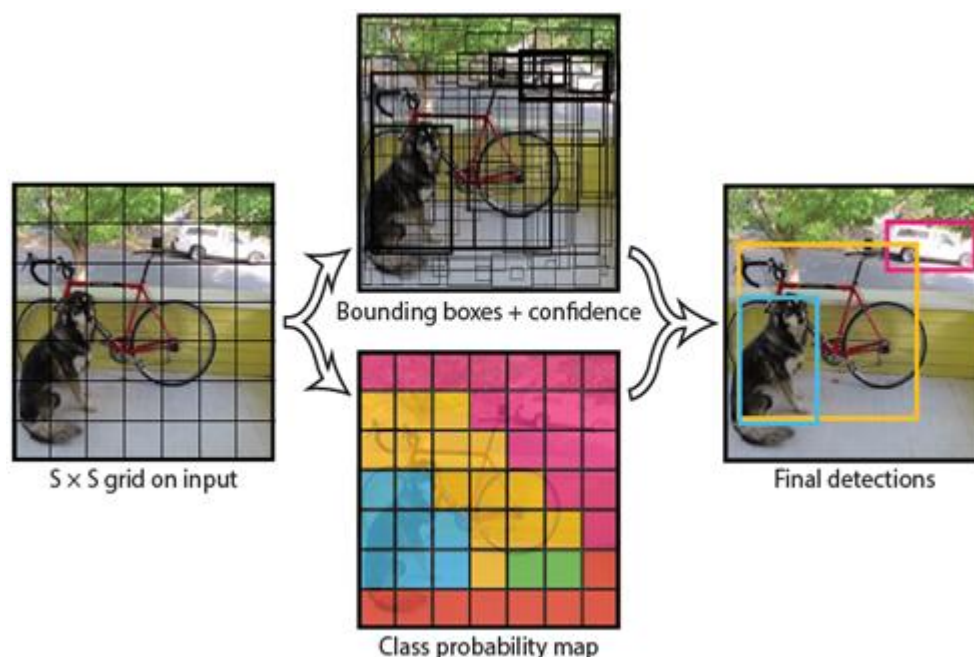
Batch Normalization normalizes a mini-batch of data in all observations for each channel independently. A Max Pooling layer performs the reduction of the resolution by dividing the input into rectangular grouping regions and by calculating the maximum of each region. A ReLU layer performs a threshold operation for each element in the input, where any value less than zero is set to zero. The YOLO v2 Transform layer extracts activations from the last convolutional layer and transforms the boundary box predictions so that they fall within the boundaries of the real box (REDMON; FARHADI, 2017).

Evaluation metrics

In problems such as object detection, the classifier's evaluation is performed by checking the correctness and possible errors of the algorithm. Hit/error possibilities can be organized and represented in a matrix known as a fuzzy matrix or contingency table.

The fuzzy matrix is made up of four variables, namely: True positives (TP) are examples correctly labeled as positive. False positives (FP) refer to negative examples incorrectly labeled as positive. True negatives (TN) correspond to negative examples correctly labeled negative, and false negatives (FN) are the number of positive samples incorrectly labeled as negative (DAVIS; GOADRICH, 2006). A fuzzy matrix is shown in Table 2.

Figure 6 - Example of several bounding boxes



Source: Redmon and Farhadi (2017).

Table 2 - Fuzzy matrix

Predicted Value			
Yes	No		
True positive (TP)	False negative (FN)	Yes	Actual
False positive (FP)	True negative (VN)	No	

Source: Davis and Goadrich (2006).

Some evaluation metrics can be associated with the fuzzy matrix, namely, precision, sensitivity, specificity, F_1 -Score, precision, and recall. These metrics were used herein to validate the interpretation of the images by the algorithm. In object detection problems, it is common to use precision, F_1 -Score, and recall as parameters for evaluation of the results (LISON; MAVROEIDIS, 2017). These metrics are defined in equations 4 to 6 below.

$$Recall = \frac{TP}{TP+FN} \quad \text{Eq. 4}$$

$$Precision = \frac{TP}{TP+FP} \quad \text{Eq. 5}$$

$$F_1 = 2x \left(\frac{Precision \times Recall}{Precision+Recall} \right) \quad \text{Eq. 6}$$

The precision measures the fraction of examples classified as positive that are truly positive in the universe, including all examples classified as positive. In turn, the recall metric measures the fraction of positive examples that are correctly labeled, that is, the rate of values classified as positive compared to how many should be classified as such. Finally, the F1-Score metric is a measure used to obtain a balance between precision and recall, for, as noted in Equation 6, it consists of the weighted mean between the first two metrics (LISON; MAVROEIDIS, 2017).

Results and discussions

The network architecture used herein is composed of 150 layers, namely: 1 input layer, 13 Addition layers, 45 Batch Normalization layers, 46 Convolutional layers, 1 Max Pooling layer, 42 ReLU layers, 1 YOLO v2 Transform layers, and 1 output layer. The algorithm used for optimization was the stochastic gradient descent with momentum optimizer. The size of the mini batches was defined as 10, the initial learning rate was 0.01, and the maximum number of epochs was 100.

Network training results

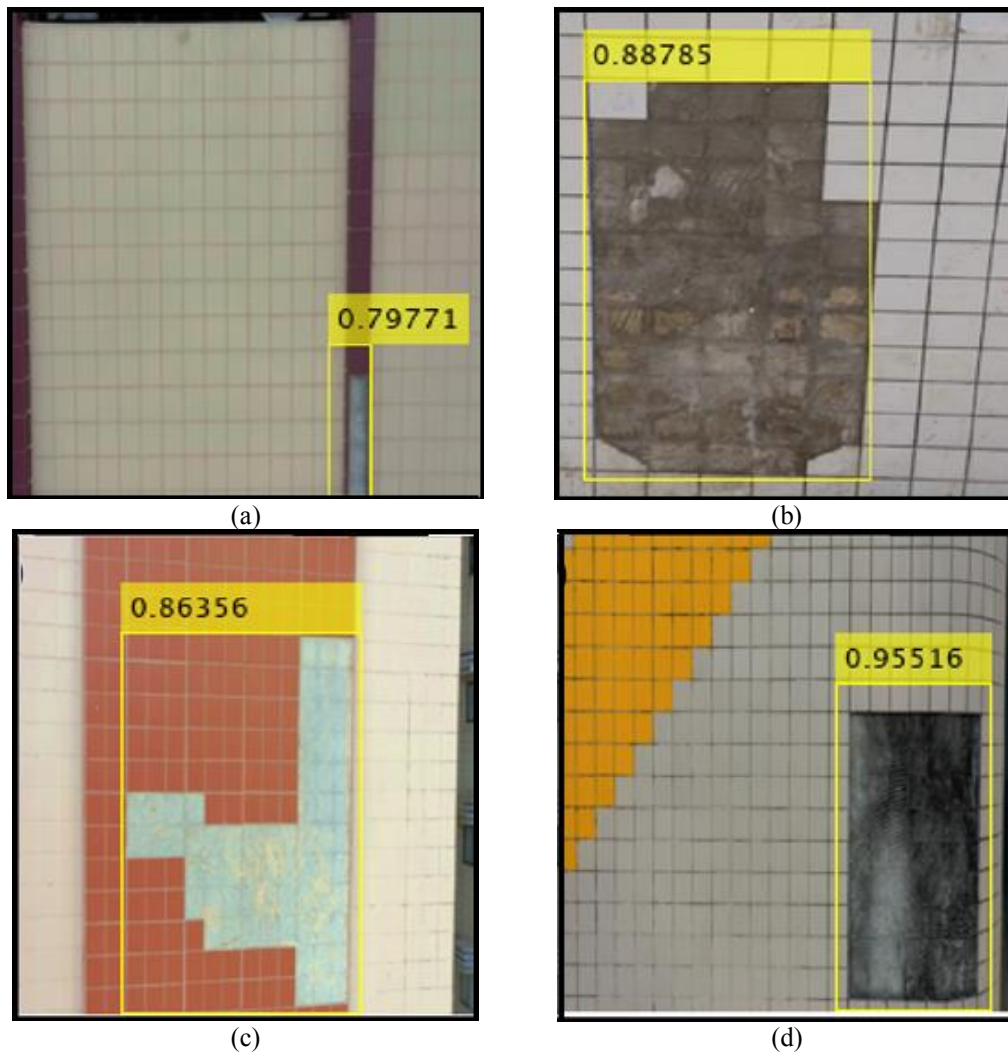
Network training was performed using 80% of the database, that is, 550 images. The reflection operation was applied to each image, doubling the number of training images, thus resulting in a total of 1100 images. This stage was performed in the Windows operating system using an Intel Core i5-9300H 2.4 GHz processor, 16 GB RAM, and 4 GB Nvidia GeForce GTX 1650 graphics processing unit. The time to carry out the training was 16 h.

Precision, recall and F_1 values were calculated based on the training data of the network in order to verify the performance of the detector on the training data. The values obtained were: precision = 0.99, recall = 0.98 and $F_1 = 0.99$. Since in this phase the images applied to the detector were the same as those used in the training, high evaluative metric values were expected. This procedure was performed only for an overview of the “behavior” of the network in the training.

Results of the testing phase in photos

The most important piece in the results section is the analysis of detector performance on the test images (20% of the total) and on the video captured with UAV. The following values were found for the test images: **precision = 0.94**, **recall = 0.75**, and **$F_1 = 0.83$** . Figures 7a - 7d show examples of how the algorithm identified areas with cladding detachment. Values above each bounding box correspond to the confidence index of the respective detection.

Figure 7 - Example of detection of areas with cladding detachment



In the results obtained in the testing phase the recall value = 0.75 is noteworthy, for it indicates a difficulty in minimizing false negative cases, that is, there were some areas with detaching cladding that went undetected. On the other hand, the high precision value means that the areas which the network identified as showing cladding detachment had in fact such pathological manifestation.

Another evaluation carried out from the fuzzy matrix was the analysis of the PR (Precision x Recall) space. In the PR space, Recall is plotted on the x axis and Precision on the y axis. This analysis consists in observing the value of the area under the curve; the closer to 1 this value is, the better the performance of the detector. Figure 8 shows the Precision x Recall graph.

The area under the curve in Figure 8 was 0.71, which demonstrates that there is still room for improvement of the detector. This is graphically visible in the considerable space in the upper right corner of the graph.

In addition to the evaluation using the classic metrics recommended in the literature, each image was visually inspected as to the results obtained in the testing phase. A finding that deserved attention in this investigation was that, in some cases, there was an overlap of bounding boxes (Figure 9).

It is worth mentioning that in the cases such as the one shown in Figure 9, the confidence index of the bounding box that best represented the pathological area was always much higher than the index of the secondary box. This indicates that the overlap problem can be solved by choosing the box with the highest confidence index.

Results of the testing phase in the video

When the detector was applied to the video recorded by the UAV, the methodology also proved to be a promising technique; however some details (described below) must be considered. Table 3 summarizes the results obtained in the analysis of the video.

Figure 8 - Recall x Precision graph

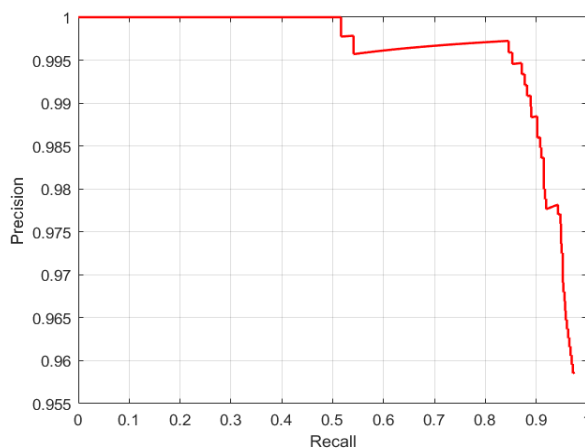


Figure 9 - Example of overlapping bounding boxes



Table 3 - Results of the application of YOLO v2 on the video record of the facade

Description	Quantity	Percentage
Area with cladding detachment	23	100%
Area identified by YOLO	21	91.3%
Unidentified area	2	8.7%

With regard to applicability in videos, two points stand out. The first point is that, in a few cases, the network failed to identify areas with cladding detachment (2 areas). This occurred mainly in images where the detached area corresponded to practically the whole frame (Figure 10).

When looking at Figure 8, establishing an optimal distance seems to be necessary so that the UAV can capture images in such a way as to avoid situations that result in photos containing areas completely devoid of cladding. Other option is to use pixel classifiers combined with object detection.

The second observation to be highlighted is that the detector “confused” two elements that constituted the facade with the occurrence of detachment. In the video used for this application, it was found that the convolutional neural network identified a guardrail as detachment (Figure 11).

These detections are the typical false positive cases and encourage the creation of methodologies that can solve this type of problem. At first, one can think about training the network with a multiclass methodology, making the detector learn to identify the different elements of the facade. Another option that can be investigated is the use of recording protocols on facades that minimize the capture of these elements.

Comparison of results between ceramic detachment and other YOLO applications

The YOLO network has stood out for its efficiency in a diversity of applications ranging from military objectives (CHEN *et al.*, 2020) to fish tracking (BARREIROS *et al.*, 2021). The application fields of YOLO include civil construction and in this area several problems have been solved by using this algorithm. Table 4 shows a comparative summary between the results obtained by using YOLO in the detection of ceramic detachment (the subject matter of this research) and other types of applications in the field of civil engineering.

By observing Table 4, it can be seen that in most applications the precision of the CNN (YOLO) being used exceeds 90%, which indicates algorithm versatility in detecting objects in an efficient way. For the application proposed hereby, the network also demonstrated precision similar to other applications (91.3% and 94%), thus demonstrating the applicability of the algorithm in the problem under analysis.

Figure 10 - Example of an area not identified by the network

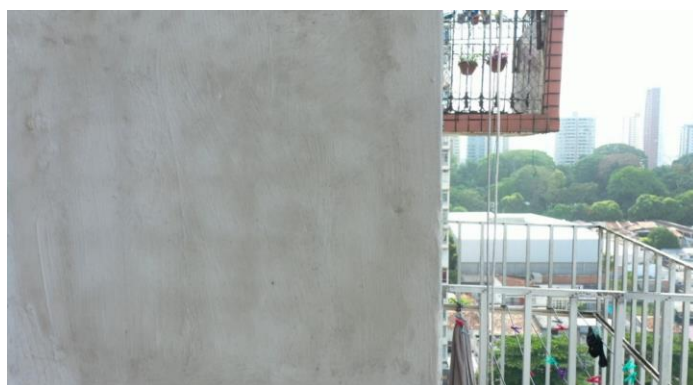


Figure 11 - Guardrail identified as cladding detachment



Table 4 - Comparison between ceramic detachment detection and other YOLO v2 applications

Authors	Object detected	Precision (%)
Hou, Zhang and Hou (2020)	Vehicles in construction	94.79
Son and Kim (2021)	Machine Operators	98.47
Zheng, Yao and Xu (2019)	Glove and Hard hat	94.79 and 94.74 (respectively)
Bo <i>et al.</i> (2019)	Hard hat	96.58
Zhang <i>et al.</i> (2020)	Damages in asphalt	91.00
Mandal, Uong and Adu-Gyamfi (2018)	Road cracks	88.51
Result of this research (photos)	Ceramic Detachment	94.00
Result of this research (video record)	Ceramic Detachment	91.30

Conclusions

This research proposes the use of computer vision to detect cladding detachment on building facades. In the results section, positive points and some specific limitations were raised, but the main objective of the research was achieved, and the cladding detachment detection model presented hereby achieved promising results. The results demonstrated the applicability of the computer vision technique YOLO v2 for automation of the inspection process of facades with ceramic detachment.

After the application of the object detection algorithm to detect cladding detachment on facades and analysis of the results, some aspects should be highlighted. First, although the image dataset was small for a deep learning network, the model performed well in the training phase. This was confirmed by the precision (0.99), recall (0.98), and F_1 (0.99) values found in this phase.

As for the testing phase, the precision in identifying areas with pathological manifestations stood out. This demonstrates that, despite some limitations, the model is applicable to the subject matter being addressed herein. The applicability of this methodology is not only feasible but also very promising and can contribute to improvements in facade maintenance management, while speeding up the inspection process.

In order to contribute to the discussion about improvements to the model introduced herein, some observed limitations are listed, namely:

- (a) overlapping bounding boxes; non-identification of some areas with the pathology (False Negative); and
- (b) the fact that the network confused facade objects with an area with no cladding (False Positive).

It is worth mentioning that, although these occurrences were minimal, they must be addressed in order to achieve an optimal model. The overlapping bounding boxes issue can be solved by choosing the box with the highest confidence index.

Overall, it can be said that this research has provided two scientific contributions to the state of the art of the topic concerned. The first consists in starting to bridge the knowledge gap presented in the introduction, that is, introducing a methodology for automated identification of cladding detachment on building facades. The second, on the other hand, consists in raising questions to be solved, so that the methodology presented herein can reach an optimal point of application.

In addition to the conclusions above, this research opens up some possibilities for future research work, including: using an image capture protocol to improve the problem of false positives and false negatives, the automatic calculation of the amount of detached ceramic tiles in order to assist in the budget for the restoration of the facade, or the use of other computer vision algorithms for the same purpose in order to compare the precision obtained.

References

- ALHEEJAWI, S. *et al.* Deep learning-based histopathological image analysis for automated detection and staging of melanoma. In: AGARWAL, B. *et al.* **Deep learning techniques for biomedical and health informatics**. New York: Academic Press, 2020.
- ALI-GOMBE, A. *et al.* Face detection with YOLO on Edge. In: ENGINEERING APPLICATIONS OF NEURAL NETWORKS CONFERENCE, 22., Halkidiki, 2021. **Proceedings [...]** Cham, 2021.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 5674**: manutenção de edificações: requisitos para o sistema de gestão de manutenção. Rio de Janeiro, 2012.
- AYOB, A. F. *et al.* Analysis of Pruned Neural Networks (MobileNetV2-YOLO v2) for Underwater Object Detection. In: NATIONAL TECHNICAL SEMINAR ON UNMANNED SYSTEM TECHNOLOGY, 11., Singapore, 2019. **Proceedings [...]** Singapore, 2021.
- BALLESTEROS, R. D.; LORDSLEEM JUNIOR, A. C. Veículos Aéreos Não Tripulados (VANT) para inspeção de manifestações patológicas em fachadas com revestimento cerâmico. **Ambiente Construído**, Porto Alegre, v. 21, n. 1, p. 119–137, jan./mar. 2021.
- BARBA-GUAMÁN, L. *et al.* Object detection in rural roads through SSD and YOLO framework. In: ROCHA, Á. *et al.* **Trends and applications in information systems and technologies**. Cham: Springer, 2021.

- BARREIROS, M. de O. *et al.* Zebrafish tracking using YOLO v2 and Kalman filter. **Scientific Reports**, v. 11, n. 1, p. 1–14, 2021.
- BAUER, E.; CASTRO, E. K.; SILVA, M. N. B. Estimate of the facades degradation with ceramic cladding: Study of Brasília buildings. **Cerâmica**, v. 61, n. 358, p. 151–159, 2015.
- BAUER, E.; MILHOMEM, P. M.; AIDAR, L. A. G. Evaluating the damage degree of cracking in facades using infrared thermography. **Journal of Civil Structural Health Monitoring**, v. 8, n. 3, p. 517–528, 2018.
- BO, Y. *et al.* Helmet detection under the power construction scene based on image analysis. In: INTERNATIONAL CONFERENCE ON COMPUTER SCIENCE AND NETWORK TECHNOLOGY, 7., Dalian, 2019. **Proceedings [...]** Dalian, 2019.
- CHEN, H.-W. *et al.* Advanced automated target recognition (ATR) and multi-target tracker (MTT) with electro-optical (EO) sensors. In: APPLICATIONS OF MACHINE LEARNING 2020. INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS, 2020. **Proceedings [...]** 2020.
- CHEN, K. *et al.* Geo-registering UAV-captured close-range images to GIS-based spatial model for building façade inspections. **Automation in Construction**, v. 122, p. 103503, 2021.
- DA COSTA, N. L.; DE LIMA, M. D.; BARBOSA, R. Evaluation of feature selection methods based on artificial neural network weights. **Expert Systems with Applications**, v. 168, p. 114312, 2021.
- DAVIS, J.; GOADRICH, M. The relationship between precision-recall and ROC curves Jesse. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 23., Pittsburgh, 2006. **Proceedings [...]** Pittsburgh, 2006.
- DIAZ, C.; CORNADÓ, C.; ALBAREDA, A. Damage in face-brick facades placed between concrete slabs. **Journal of Building Engineering**, v. 30, p. 101312, 2020.
- GALANTUCCI, R. A.; FATIGUSO, F. Advanced damage detection techniques in historical buildings using digital photogrammetry and 3D surface analysis. **Journal of Cultural Heritage**, v. 36, p. 51–62, 2019.
- HOU, X.; ZHANG, Y.; HOU, J. Application of YOLO V2 in Construction Vehicle Detection. In: MENG, H. *et al.* Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery. Cham: Springer, 2020.
- IOFFE, S.; SZEGEDY, C. **Batch normalization**: accelerating deep network training by reducing internal covariate shift. 2015. Available: <https://arxiv.org/pdf/1502.03167.pdf>. Access: 07 jan. 2021.
- ISLAM, M. M. M.; KIM, J.-M. Vision-based autonomous crack detection of concrete structures using a fully convolutional encoder–decoder network. **Sensors**, v. 19, n. 19, p. 1–12, 2019.
- JEONG, G. Y. *et al.* Applying unmanned aerial vehicle photogrammetry for measuring dimension of structural elements in traditional timber building. **Measurement**, v. 153, p. 107386, 2020.
- KATTENBORN, T. *et al.* Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 173, p. 24–49, 2021.
- KAVUMA, A.; OCK, J.; JANG, H. Factors influencing Time and Cost Overruns on Freeform Construction Projects. **KSCE Journal of Civil Engineering**, v. 23, n. 4, p. 1442–1450, 2019.
- KERLE, N. *et al.* UAV-based structural damage mapping: a review. **ISPRS International Journal of Geo-Information**, v. 9, n. 1, p. 1–23, 2019.
- LI, S.; ZHAO, X. Image-based concrete crack detection using convolutional neural network and exhaustive search technique. **Advances in Civil Engineering**, v. 2019, n. MI, 2019.
- LISON, P.; MAVROEIDIS, V. **Automatic detection of malware-generated domains with recurrent neural models**. 2017. Available: <https://arxiv.org/pdf/1709.07102.pdf>. Access: 07 jan. 2021.
- LOEY, M. *et al.* Fighting against COVID-19: a novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. **Sustainable Cities and Society**, v. 65, p. 102600, 2021.
- MADUREIRA, S. *et al.* Maintenance planning of facades in current buildings. **Construction and Building Materials**, v. 147, p. 790–802, 2017.
- MANDAL, V.; UONG, L.; ADU-GYAMFI, Y. Automated road crack detection using deep convolutional neural networks. In: 2018 IEEE INTERNATIONAL CONFERENCE ON BIG DATA (BIG DATA), Seattle, 2018. **Proceedings [...]** Seattle, 2018.

- MATHWORKS. **Get started with the image labeler**. MathWorks. Available: <https://www.mathworks.com/help/vision/ug/get-started-with-the-image-labeler.html>. Access: 20 jun. 2020a.
- MATHWORKS. **Getting Started with YOLO v2**. Available: <https://www.mathworks.com/help/vision/ug/getting-started-with-yolo-v2.html>. Access: 16 maio 2020.
- MORILLAS, H. *et al.* Nature and origin of white efflorescence on bricks, artificial stones, and joint mortars of modern houses evaluated by portable Raman spectroscopy and laboratory analyses. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 136, n. PB, p. 1195–1203, 2015.
- PAN, Y.; ZHANG, L. Roles of artificial intelligence in construction engineering and management: a critical review and future trends. **Automation in Construction**, v. 122, n. August 2020, p. 103517, 2021.
- PEREIRA, F. C.; PEREIRA, C. E. Embedded image processing systems for automatic recognition of cracks using UAVs. **IFAC-PapersOnLine**, v. 28, n. 10, p. 16–21, 2015.
- PEREZ, H.; TAH, J. HM; MOSAVI, A. Deep learning for detecting building defects using convolutional neural networks. **Sensors**, v. 19, n. 16, p. 3556, 2019.
- PIRES, R.; DE BRITO, J.; AMARO, B. Inspection, diagnosis, and rehabilitation system of painted rendered façades. **Journal of Performance of Constructed Facilities**, v. 29, n. 2, p. 04014062, 2015.
- PLASTIRAS, G.; KYRKOU, C.; THEOCHARIDES, T. **Efficient convnet-based object detection for unmanned aerial vehicles by selective tile processing**. 2019. Available: <https://arxiv.org/pdf/1911.06073.pdf>. Access: 07 jan. 2022.
- REDMON, J.; FARHADI, A. YOLO9000: better, faster, stronger. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 30., Honolulu, 2017. **Proceedings [...]** Honolulu, 2017.
- RUIZ, R. D. B. *et al.* Processamento digital de imagens para detecção automática de fissuras em revestimentos cerâmicos de edifícios. **Ambiente Construído**, Porto Alegre, v. 21, n. 1, p. 139–147, 2021.
- SILVESTRE, J. D.; DE BRITO, J. Ceramic tiling in building façades: Inspection and pathological characterization using an expert system. **Construction and Building Materials**, v. 25, n. 4, p. 1560–1571, 2011.
- SINGH, S. A.; MAJUMDER, S. Short and noisy electrocardiogram classification based on deep learning. In: SINGH, S. A.; MAJUMDER, S. **Deep learning for data analytics**. New York: Academic Press, 2020.
- SON, H.; KIM, C. Integrated worker detection and tracking for the safe operation of construction machinery. **Automation in Construction**, v. 126, p. 103670, 2021.
- SONG, W. *et al.* Automatic pavement crack detection and classification using multiscale feature attention network. **IEEE Access**, v. 7, p. 171001–171012, nov. 2019.
- SREENATH, S. *et al.* Assessment and use of unmanned aerial vehicle for civil structural health monitoring. **Procedia Computer Science**, v. 170, p. 656–663, 2020.
- VALERO, E. *et al.* Automated defect detection and classification in ashlar masonry walls using machine learning. **Automation in Construction**, v. 106, p. 102846, jun. 2019.
- YOSINSKI, J. *et al.* How transferable are features in deep neural networks? **Advances in Neural Information Processing Systems**, v. 4, p. 3320–3328, jan. 2014.
- YU, Y. *et al.* A novel deep learning-based method for damage identification of smart building structures. **Structural Health Monitoring**, v. 18, n. 1, p. 143–163, 2019.
- ZHANG, C. *et al.* Construction worker hardhat-wearing detection based on an improved BiFPN. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 25., 2020. **Proceedings [...]** 2021.
- ZHANG, J. *et al.* Automatic detection of moisture damages in asphalt pavements from GPR data with deep CNN and IRS method. **Automation in Construction**, v. 113, p. 103119, 2020.
- ZHENG, X.; YAO, J.; XU, X. Violation monitoring system for power construction site. **IOP Conference Series: Earth and Environmental Science**, v. 234, article 012062, 2019.

Acknowledgements

The authors would like to acknowledge the support of the Coordination for the Improvement of Higher Education Personnel (CAPES) - Brazil.

Alcineide Dutra Pessoa de Sousa

Programa de Pós-Graduação em Engenharia Civil | Universidade Federal do Pará | Rua Augusto Corrêa, 01, Guamá | Belém - PA - Brasil | CEP 66075-110 | Tel.: (91) 3201-8859 | E-mail: alcineidedutra@hotmail.com

Gean Carlos Lopes de Sousa

Bacharelado em Engenharia de Alimentos | Universidade Federal do Maranhão | Av. da Universidade, s/n, Bom Jesus | Imperatriz - MA - Brasil | CEP 65915-240 | Tel.: (99) 3259-6055 | E-mail: gean.sousa@ufma.br

Luiz Maurício Furtado Maués

Programa de Pós-Graduação em Engenharia Civil | Universidade Federal do Pará | E-mail: maues@ufpa.br

Ambiente Construído

Revista da Associação Nacional de Tecnologia do Ambiente Construído

Av. Osvaldo Aranha, 99 - 3º andar, Centro

Porto Alegre - RS - Brasil

CEP 90035-190

Telefone: +55 (51) 3308-4084

www.seer.ufrgs.br/ambienteconstruido

www.scielo.br/ac

E-mail: ambienteconstruido@ufrgs.br



This is an open-access article distributed under the terms of the Creative Commons Attribution License.