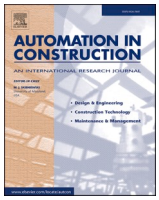
Automation in Construction 145 (2023) 104658

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon



Deep learning model for automated detection of efflorescence and its possible treatment in images of brick facades

David Marín-García [a,*], David Bienvenido-Huertas [b], Manuel J. Carretero-Ayuso [c], Stefano Della Torre [d]

[a] Department of Graphical Expression and Building Engineering, Higher Technical School of Building Engineering, University de Seville, Seville 41012, Spain
[b] Department of Building Construction, University of Granada, Granada 18001, Spain
[c] Department of Architecture, University of Alcalá, Alcalá de Henares 28801, Spain
[d] Department of Architecture, Built Environment and Construction Engineering, Politecnico di Milano, Italy

ARTICLE INFO

Keywords:
Computer vision
Deep learning
Repair of efflorescence from brick facades

ABSTRACT

One of the most common pathologies in exposed brick facades is efflorescence, which, although they often have a similar appearance, their effects and way of solving them can range from a one-off cleaning to a repair that involves adding or replacing the material. Therefore, the novel **goal** of this work is to verify whether it is possible to automate this task of distinguishing what type of intervention each brick needs. To do this, the **methodology** followed focuses on proposing, training and validating a deep convolutional neural network with the real-time end-to-end method that simultaneously predicts multiple bounding boxes and class probabilities for those boxes. For this, images of 765 building facades will be used, of which 392 were selected, proceeding to label 4704 bricks, **resulting** in that the model achieved a mAP maximum at epoch 100 with 0.894, which is therefore of interest for the creation of intervention maps.

1. Introduction

Deep learning [1–3] based on convolutional neural networks has been frequently used to solve image classification problems [4] both generic [5–10] and specific to medicine [11–13], industry [14], environment [15], among many others [16]. Of all of them, due to the theme of this study, it is necessary to highlight the applications made in the field of construction [17], its processes [18] and the architectural heritage [19,20]. One of the applications in this field consists in identifying various pathologies based on the symptoms reflected in images obtained through different methodologies, including mobile devices, information and communication technologies, and existing image databases. Thus, some studies conducted on building have ordered these pathologies and their symptoms according to the frequency with which they appear [21], and there are investigations related to the application of deep learning to identify the most frequent ones, such as fissures and cracks [22], mold, stain, and paint deterioration [23], including peeling, blistering, flacking, and crazing, among others. However, artificial vision must continue advancing in this field, and that is why this research focuses on the training of a model that allows detecting one of the most frequent damages in buildings around the world, whose facades are made with bricks ceramics, exposing their faces. These damages are called efflorescence and are manifested by the appearance of generally whitish surface salts. In this sense, some works such as the work carried out by authors such as Wang et al. and its "Automatic damage detection in historic masonry buildings based on mobile deep learning" [24] are very interesting. However, this kind of work, although among other pathologies, detects efflorescence, is carried out on the masonry of historic buildings and the type of intervention is not classified. Thus, in the present work, the use of artificial vision is not only intended to locate bricks affected by efflorescence. It is intended to go even further, checking how a properly trained model with a low-density data set can distinguish in real-time subtleties such as which and how many bricks have not deteriorated due to the efflorescence said and which have, even classifying those that need cleaning simple and those that would require a greater intervention due fundamentally to the loss of material. Consequently, the objective of this study is based on verifying the results of the application of an artificial vision tool to help in exclusively visual inspections carried out by experts.

* Corresponding author.
E-mail address: damar@us.es (D. Marín-García).

https://doi.org/10.1016/j.autcon.2022.104658
Received 12 August 2022; Received in revised form 30 October 2022; Accepted 31 October 2022

1.1. Background

1.1.1. Deep learning based on convolutional neural networks used to solve image classification problems

Regarding the so-called deep learning, we must remember that it is initially based on the biological study of learning in the 1940s [25,26], the appearance of the first perceptron models in the 1950s [27] and the techniques of backpropagation [28], in the 1980s, which is currently one of the most used techniques in artificial neuron network (ANN). On the other hand, if we have an ANN with many hidden layers, we are in the presence of a deep network. However, the concept of deep learning goes further and is difficult to define, although in the case of algorithms, it is understood that we are in the presence of deep learning when a high number of transformations of the input signal is carried out with respect to the one that comes out.

On the other hand, LeCun in 1998 [29] carried out the first experiments on the use of a convolutional algorithm applied to neural networks for image recognition, applying filters (kernel) to the data that make up the images, which allow them to be simplified by identifying important patterns or characteristics. The network for recognizing images of numbers was called LeNet, which had four convolutional hidden layers plus as many final no-convolutional layers. In convolutional networks, the neurons basically learn the Kernerl, that is, these will be their weights and instead of being connected to all the neurons of the previous layer, what it does is convolve on the previous layer. The output is called a feature map. There are numerous kernel arrays that are also used in graphics programs to highlight outlines, blur, etc. The model obtained is completely derivable in order to carry out gradient propagation. Between convolutional layers, Pooling is also usually introduced from time to time, which allows reducing the resulting matrices. There are several types (MAX, AVG, SUN, L2, etc.) For example, the MAX which keeps the maximum values of various values.

Since then, based on these principles of deep learning associated with convolutional neural networks (CNN), numerous algorithms have been developed and applied in many environments [30–32]. There have even been global competitions where these algorithms and architectures are tested on a large number of images, such as ImageNet and Large Scale Visual Recognition Challenge (ILSVEC) [33].

Regarding the way in which these automatic image classification methods are usually materialized, first, it should be remembered that it is usually supervised learning, that is, a suitably labeled image data set is needed to train and validate these algorithms.

As for what algorithms are available, which one to choose and how to implement it, we can say that there is no single solution. Usually, deep learning frameworks or preconfigured platforms are used. The former, such as Tensorflow, Keras or PyTorch, among others, are usually open source and require certain programming knowledge (Python, C, Java, etc.) to apply them, as well as a computer with a powerful GPU and other specific features, although cloud services, such as online GPU instances in Google Colab, can also be used, but with certain limitations. Another option is pre-configured API platforms. However, the biggest disadvantage of the latter is that in many cases they require payment, at least for large amounts of data and processing time, and although they are very powerful tools, they do not give as much freedom to the developer.

In the specific case that concerns us (detecting bricks with efflorescence that are contained in an image of a building facade), it is about recognizing objects within an image and distinguishing the different classes based on their characteristics. To accomplish this task, there are deep learning-based object detection models that can classify the object and detect the position of this in the image by drawing a bounding box around it. Some can even generate segmentation masks (Mask RCNN) that perfectly reflect the object for each instance [34].

This type of object detection has been developed mainly since the beginning of the 21st century [35]. These first algorithms used handcrafted features, but since 2010 the need for advance to improve the results has been detected. Thus, a few years later, R. Girshick uses the so-called Regions [5] together with the CNN functions, thus appearing the R-CNN that presented results that duplicated those previously offered. However, the drawback was still the detection time, especially for real-time images. For this reason, other proposals emerged, such as that of K. He et al. and its proposed Spatial Pyramid Pooling Networks (SPPNet) [36] or the Spatial Pyramid Pooling (SPP) layer, which allows a convolutional neural network to develop a fixed length representation independent of the size of the image/region of interest and does not require change of scale. This innovation allowed greater speed without loss of precision.

However, only a year later, one of the most important advances occurs, the one made by R. Girshick and his Fast R-CNN detector [37], which is a couple of hundred times faster than the previous R-CNN. Also noteworthy is the Faster R-CNN by S. Ren et al. [38], which included a Region Proposal Network, and the R-FCN [39], which included position-sensitive score maps for faster detection.

But it is J. Redmon et al. the one that makes one of the most outstanding contributions to date in 2015 and is the so-called YOLO (You Only Look Once) [40] that does not use regions to locate objects in the image, detecting with high precision (sometimes close to or even greater than 90%) and in milliseconds, multiple objects after observing the complete image. Also noteworthy in this regard is the contribution of W. Liu et al. with its SSD (Single Shot Multiple Box Detector) [41] that, although it offers great precision, like other options such as RetinaNet [42], they are not as fast as the latest versions of Yolo [43].

In summary, it can be said that there are two types of visual object detection. The first is the region proposal method (e.g. R-CNN, Mask RCNN, Fast R-CNN, Faster R-CNN, etc.), and the second is the end-to-end method (e.g. Yolo, SSD, RetinaNet, etc.).

As object detection technology has evolved, the Yolo series of algorithms, with very high precision and speed, has been used in various scene detection tasks. At the same time, the system computes all the image's features and predicts all the objects. Version v5 is the fifth generation, written in Python programming language, and according to various studies, outperforms the rest of the models in terms of accuracy and speed [44]. To the best of our knowledge, this is the first time it has been used to detect the type of treatment of facing bricks on building facades affected by efflorescence in the present study.

1.1.2. Efflorescence in exposed ceramic bricks for building facades

Exposed brick facades and their pathologies and defects [45], whether aesthetic or related to their function as a building envelope, have been the subject of numerous studies [46]. Thus, one of the most frequent conditions that appear in these elements is the efflorescence.

In addition, efflorescence in buildings has already been studied both in terms of its classification [47–49] and in terms of its prevention and elimination [50]. Therefore, in exposed ceramic brick facades, these efflorescence frequently manifest themselves by forming salts on the surface whose crystals sometimes have the shape of flowers that precisely give it its name [51] and if they form inside the material, they are likely to damage it [52]. The composition of these salts may be diverse, such as calcium carbonate or sodium sulfates, calcium, magnesium, potassium, etc., [48] and the way to solve it is different depending on the said composition, the porosity of the brick, the presence of other concomitant deterioration factors and the degree of deterioration already suffered.

Thus, if the brick is not physically damaged and no concurrent deterioration factor is detected, the solution usually consists of adequate cleaning following a certain methodology and application of specific products, as indicated in the study carried out by Rincón et al. [50]. On the other hand, if other deterioration factors are detected, or the brick has already deteriorated, the solution consists in either repairing it with specific products that restore the deteriorated material, or even completely replacing it. In these cases, more investigation will be required, punctually focusing on the bricks whose condition is not clear.

Therefore, it is evident that making a location (mapping) and prior

diagnosis of which bricks need a simple cleaning and which ones need a deeper investigation and repair is a matter of interest in interventions of this type since it can help significantly to fine-tune the repair diagnosis and to optimize the budget and planning of the work to be carried out. It is noteworthy that detailed mapping is practically useful as such interventions are expensive, and the cost calculation uses to be detailed by small quantities.

For all these reasons, it is of great interest to have an automatic system to help locate and quantify the bricks affected by efflorescence, as well as to know what type of solution each of them requires (simple cleaning or more detailed investigation and repair). In this sense, in Table 1 we can observe the basic visual aspects of efflorescence in exposed bricks that allow us to differentiate whether the method to solve them is superficial (cleaning and, where appropriate, application of suitable products) or requires a repair of the material (repair).

Although Table 1 offers a basic initial orientation for the task of labeling bricks with efflorescence and type of repair, this could generate doubts about the type and probable severity of pathological damage and its relationship with the kind of repair and as this is subjectively justified. In this sense, it must be said that it is possible to carry out "in situ" measurements with reagents and even quantitative measurements in the laboratory by X-ray diffraction, for example, to find the chemical compounds present and thus deduce the type of efflorescence and its possible consequences and even ways to solve it. However, given the high number of facades used, except in some specific verification cases, it was not possible to perform these tests and therefore, it may be an issue to be addressed in future research to find out if this can improve the results. Despite this, it must be said that very useful organoleptic tests were carried out to know, for example, whether or not there is disintegration of the material, closely related to the need to go beyond a simple cleaning in the intervention, even in the early stages of pathological development. These tests and the experience of experts have been the basis for the necessary labeling.

1.2. Deep learning applied to the classification of the efflorescence resolution method in facing bricks in building facades from their images

As mentioned above, several studies have been conducted in relation to the detection by computer vision of numerous pathologies in buildings and constructions of various types [17,22,23]. Some even address the issue of efflorescence detection [24].

However, no specific studies related to the application of deep learning to image classification have been detected to help map efflorescence in exposed brick facades in buildings in terms of their classification depending on the next steps to be applied to obtain a solution to the problem.

Table 1
Basic visual aspects of efflorescence in bricks and the method to solve it.

Usual visual aspect	Usual consequences according to references-experts	Usual solution
Surface deposit of whitish salts in the central area of the brick, without deterioration of the brick.	Aesthetic damage. In general, it does not cause significant deterioration of the material.	**Clean** (cleaning and, where appropriate, application of suitable products).
Chipped pieces and degradation of the surface layer to a greater or lesser extent that is easily detached, with whitish salts appearing inside.	Material and cosmetic damage. They can affect the durability of the brick.	**Repair** (requires a deeper investigation of the material and usually substitution).
Whitish deposits from the cement of the mortar.	In general, there is no significant deterioration in the material. .	They will **not be the object of this study.**

2. Materials and methods

This section explains the equipment and materials to be used, as well as the training and testing models. For the purposes of the test, the description of the state of repair of the bricks has been simplified into the binary alternative of not damaged (just to be cleaned)/damaged (to be repaired). Fig. 1 shows the process diagram.

2.1. Preparation work

Building a primary data set by taking photographs is a difficult and time-consuming task. In this case, 765 photographs of exposed ceramic brick façades of buildings with efflorescence have been taken in buildings located in Spain, specifically in Andalusia, where, although it has been used quite frequently, its use is becoming more common due to its durability and low maintenance, in addition to its acoustic and aesthetic properties. Regarding the characteristics of the bricks of the chosen facades, although it is not possible to cover all the typologies that can be presented, it has been tried to be as varied as possible within the most used brick standard in Spain and one could even say that in a large part of the countries worldwide. Thus, the chosen facades have bricks with a face or faces with a smooth visible surface, although they may have occasional imperfections or slight roughness. On the other hand, painted facades have been excluded. The geometry of the bricks of the facades that make up the dataset is rectangular, but with diversity in terms of dimensions, natural color of the brick, the type of rigs and the apart or sores between bricks. It can be thought that, with these characteristics, the use that can be made of the final algorithm is very limited, however, the defined standard is very frequent and therefore it is understood to be very useful, being the subject of another investigation the application of this methodology to another type of standards.

Table 2 provides a summary of the conditions and the characteristics of the façade materials used to form the data set.

Fig. 2 offers a sample of the photographs of the obtained dataset.

Once all images were manually reviewed, 392 were selected that met the previously established requirements of being exposed and smooth ceramic bricks and that the images did not have photographic defects that made them unfeasible. In each of these 392 images, the bricks with efflorescence or damaged by them were marked-labeled, differentiated between those that needed a simple cleaning (clean) and those that needed a major repair due to loss of material (repair). A total of 4704 smooth exposed bricks related to efflorescence and crypto- efflorescence were tagged.

To label the images, the labeling tool was used to draw the bounding box and label the class. After successfully labelling the images, it gets the output as a text file and a class file. The classes of our project are ćleań and ŕepaiŕ.

Labeling is done for both test and training data sets. The percentage of images taken for training was 80% and that for testing was 20%. This was done for each class.

2.2. Algorithm used and training with the labeled images

Once the data sets are obtained, the classifier algorithm is used. In this sense, it must be considered that what is intended is a real-time classification and location of the objects (bricks with efflorescence) within an image in such a way that differences can be distinguished as to how to intervene in said bricks to restore them. As mentioned in the Introduction, Yolov5 [55] was chosen because it offers the best trade-off between speed and accuracy (especially speed). It is an algorithm with four versions s, m, l, x, with s being the one with the highest speed and the lowest average precision, and x the one with the lowest speed and the highest precision. In this work, s and l were tested, although some verification tests with the other versions were also carried out. Regarding the size of the model, it turned out to be one of the smallest that exists within the range of object detection algorithms with faster

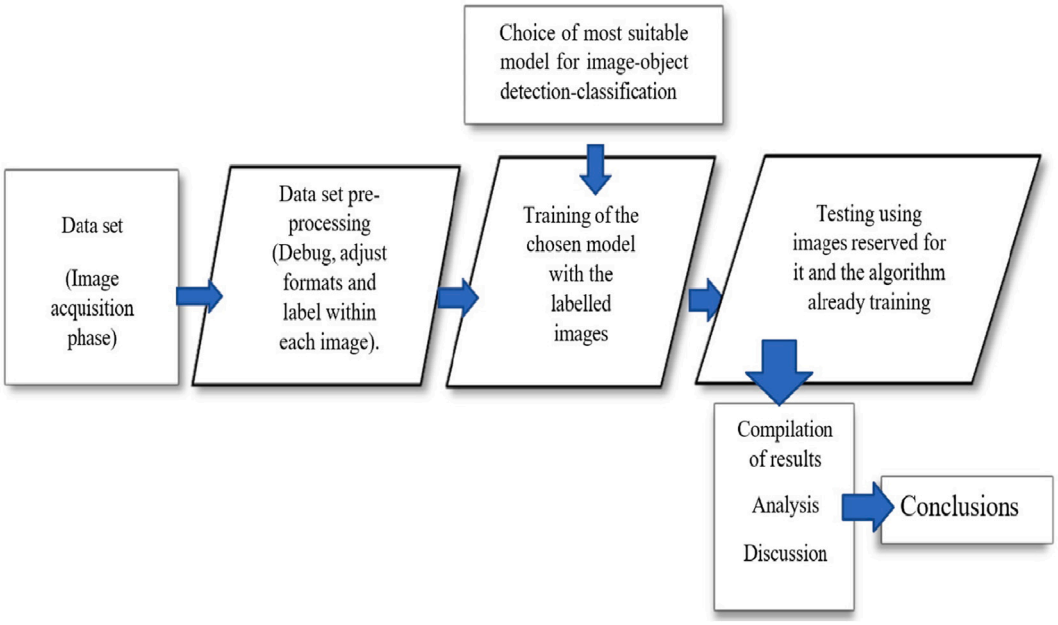


Fig. 1. Steps followed in the research.

Table 2
Summary of the conditions and the characteristics that the bricks of the facades used to make up the data set of images had to meet (based on normative standards that have been in force in Spain at different times [53,54]).

Conditioning factors	Characteristics
General original characteristics	The bricks will present regularity of dimensions and shape that allow the obtaining of mortar layers of uniform thickness, equal rows, regular facings and uniform seating.
Flatness of the faces	There will be no flection, sag or deformation on the edges and diagonals of the faces (up to 4 mm in the longest bricks and 1 mm in the shortest).
Physical characteristics	Non-freezing and compact, unpainted and with a uniform natural coloration, although they may present slight variations in tones and intensity as long as a homogeneous intonation is maintained on the façade,
Original defects not related to or caused by efflorescence	Walls that do not present generalized defects in their bricks (more than 20%) related to cracks, breaks, etc. or other conspicuous defects not related to or caused by efflorescence.

positioning and recognition speeds.

The algorithm mainly contains three parts: the spine, the neck, and the head, which are, respectively, involved in information feature extraction, feature pyramid construction, bounding box creation, and confidence score + probability map. Class, and finally the prediction bounding box and object class and the confidence score. Fig. 3 shows the structure of the model.

The process starts with the algorithm normalizing the input image to a fixed standard size of 608 × 608. The backbone includes the pre-training and the head predicts the classes and bounding boxes in a single step called dense prediction done by the already mentioned end-on-end algorithms like Yolo, SSD, RetiaNet, etc., as opposed to two-stage ones like Faster R- CNN). The layers that make up the neck collect feature maps.

3. Results and analysis

In our training, it is kept the batch size 4 and ran the training session for 20, 40, 60, 80, 100 and 120 epochs. The dataset was split into an 80:20 ratio meaning 80% of the entire dataset went to the training part, and the rest 20% went to validation. We trained our dataset using transfer learning on small and large variations of the model where the models were pre-trained on the COCO dataset. We used SGD optimiser with a learning rate of 0.01 and a weight decay of 0.0005. At epoch 40, the model achieved mAP greater than 0.80, filled to its maximum at epoch 100 with mAP of 0.894, the results varying a few tenths between vs (small version) and vl (large version) (Table 3). The large model took three times as long to train. As the training finishes, the model saves both the last and the best weights. Here, the mAP the model shows after each epoch is calculated based on the validation set.

Table 4 showcases the precision, recall and mAP scores that it is achieved using the version v5l (large model). From Table 3 it can see that the model achieved mAP score of 0.521 on the first ten epochs. Then, in epoch 10–19, the mAP score increased considerably (from 0.521 to 0.711). Later, in the epochs 20–29, 30–39, 40–49, 50–59 and 60–69, there were also significant increases, but not as much as in the first (increases of around 0.03). In the 60–69 epoch, the model also achieved a mAP score of 0.892, which together with the 90–99 epoch is the highest score the model achieved during training. After that, the mAP decreased, although not very significantly (except in the aforementioned epoch 90–99 that practically equalized).

The confusion matrix (Fig. 4) provides a deeper understanding of how our model perform in each class. The X-axis depicts the real values, and the Y-axis depicts the "Predicted" values. Thus, in the case of the detection of bricks to be cleaned (clean), we can see that the cell illustrates 0.85. This means that our model was able to correctly predict 85% of the images, which were images of bricks that needed simple cleaning. In the same way, it happens with the bricks that need repair (i.e. further investigation) being the prediction of 0.82 (82%).

Regarding the confidence vs precision graph (Fig. 5) it can be see that the graph is upward sloping. This means the average precision level is increasing against confidence.

In contrast, the confidence vs. recall graph (recall curve-Fig. 6) is negatively sloped against confidence.

Of the two classes of tags, the "clean" ones obtained the best results, then the "repair" ones, although the difference was not very significant in most cases. This happened because it is somewhat easier to detect simple cleaning when there are no signs of deterioration, especially in the presence of efflorescence that has a white color that is easily

Fig. 2. Sample of the photographs of the initially obtained data set.

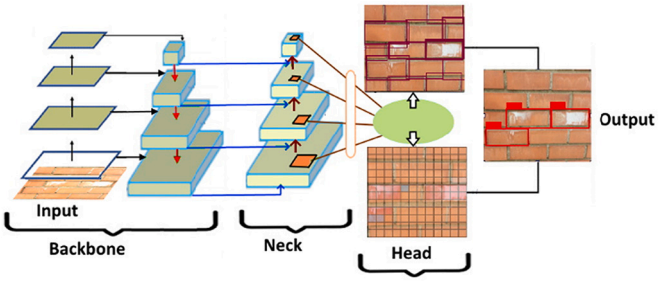


Fig. 3. Model.

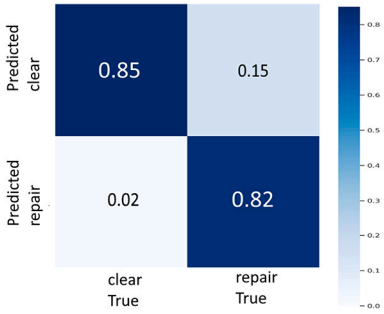


Fig. 4. Confusion matrix.

Table 3
Training time and best mAP version Small Vs Large Model.

Model	Training time	Best mAP
V5s.	The V5l took three times as long to train.	0.851
V5l		0.894

Table 4
YOLOv5l (large metrics) for 100 epochs.

Epoch	Precision	Recall	mAP
0–9	0.501	0.58	0.521
10–19	0.66	0.698	0.711
20–29	0.67	0.775	0.77
30–39	0.805	0.75	0.818
40–49	0,838	0.769	0.837
50–59	0.869	0.781	0.86
60–69	0.914	0.793	0.892
70–79	0.889	0.815	0.881
80–89	0.916	0.837	0.889
90–99	0.943	0.814	0.893

distinguishable from the brick color and with smooth textures. On the other hand, when the brick is highly degraded, it is also more significantly distinguished from the other bricks, which led the model to detect them with greater precision. On the other hand, it was more difficult for the model to detect some bricks with little deterioration, especially in its differentiation between classifying them as "clean" or "repair" since their shapes and textures are more confusing to the algorithm. Fig. 7 shows some of the predictions made by the model in the validation set.

Fig. 8 shows a test carried out on a part of a facade to visually check if the algorithm responds correctly to the requested classification, especially in the cases of "repair" in which the intervention is more expensive. In the said figure, it is also of great interest to observe the figures reached for each label, in such a way that the higher the figure, the greater the confidence that the classification is correct. In this case, the most deteriorated correspond to those with the highest figures, so a certain correspondence could be made regarding the higher or lower cost of repair. In addition, in the case of larger facades, they also automatically supply us with the mapping and number of affected bricks that must be cleaned or/and repaired.

As can also be seen in Fig. 5, some bricks have other types of alterations (cracks, cement stains, etc.) that are not classified since they are not the type of deterioration that is intended to be detected. On the other hand, it can be verified that some bricks that could be labeled for cleaning are not, however, they are a minimum quantity and, in this sense, from a future perspective, a safety coefficient could be established when calculating the magnitude of the intervention.

It must be said that after carrying out several tests with images,

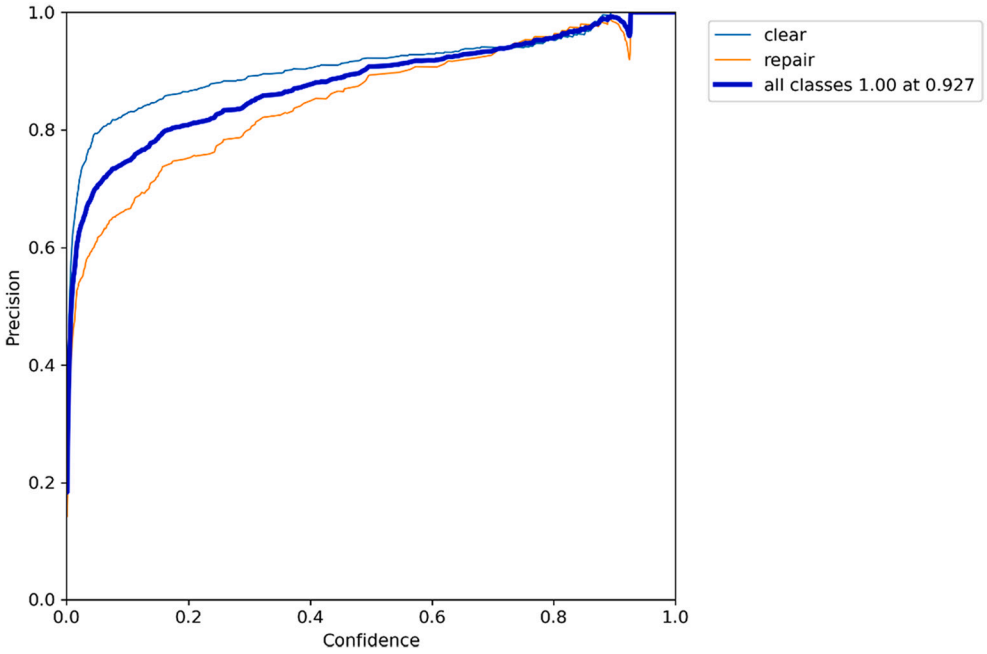


Fig. 5. Confidence vs. precision.

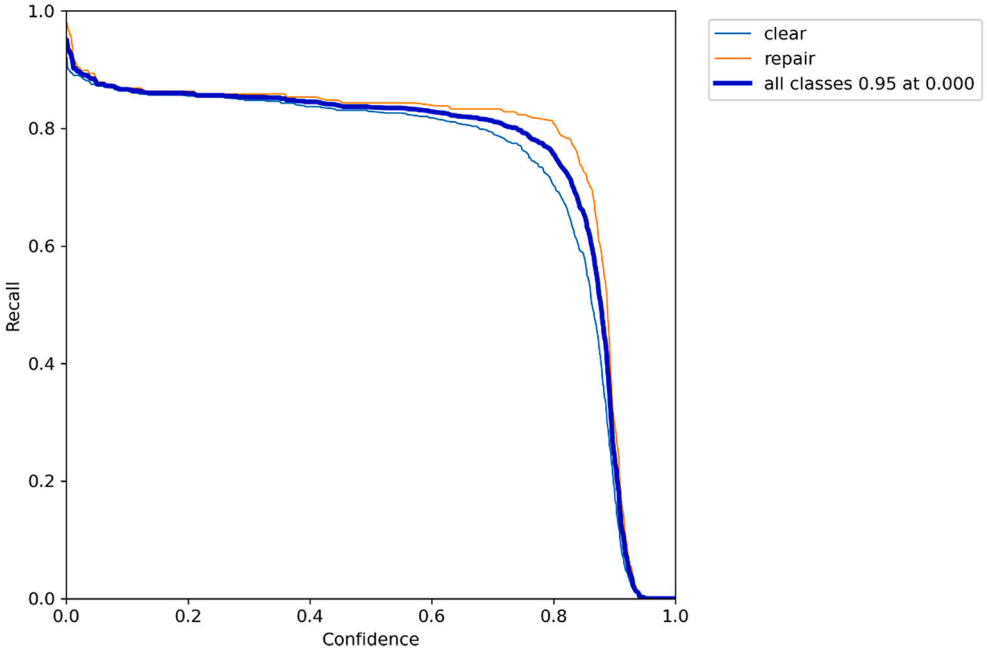


Fig. 6. Confidence vs. recall.

submitting them to the algorithm already trained, it is verified that in terms of recognition speed, 0.7 ms are achieved by pre-processing, and 17.6 ms of inference, 1.7 ms NMS per image.

Finally, it should be noted that the tests in Figs. 7 and 8 could be thought to show bricks of a similar nature, and that this would reduce the actual application of the methodology. However, as previously mentioned, it is a pilot experience that aims to be limited to facades made of frequently used bricks, widely used in several countries around the world, such as Spain, and in any case the research can be extended in the future to other types of bricks and facades.

4. Conclusions

Given that efflorescence on brick facades of buildings is one of the most frequent pathologies in construction, and that with the proposed trained model, 89% is achieved in terms of mAP@0.5 and a recognition speed of 0.7 ms per preprocess, 17.6 ms inference and 1.7 ms NMS per image, it can be said that these technologies can be of great help to create intervention maps on exposed brick facades in real time, from a set of data of low density, so it is understood that the training of the algorithm can be used in a generalized way and even in real time as an aid to mapping in the repair of this type of facades and therefore as support in the necessary interventions of repair and cleaning of

Fig. 7. Validation set prediction examples.

Fig. 8. Test to visually check if the algorithm responds correctly, especially in "repair".

buildings both newly built and belonging to the architectural heritage, thus allowing a better location of the damage and the economic and temporal quantification of the possible intervention. On the other hand, the obtained maps depend on the first diagnose input, therefore, the implemented technology can speed up mapping, but cannot substitute the diagnostic skills of technicians.

Given that it is a pilot test and that it has been carried out with a low-density data set, it is very likely that there is room for improvement in accuracy, so in the future, it is intended to continue feeding the set of data. Image data to continue training the algorithm for said improvement since the construction of the image data set that we have carried out, although it has been tried to be both varied and balanced and has shown high precision, must be extended with a greater variety of facades, bricks and forms. Future work may expand the image classes and the volume assigned to each class in the image dataset. Another limitation on which this work should continue to expand in the future is the circumstance that surely, in facades with other characteristics of the bricks, exposed to directed rain or sheltered facades, the reliability of the identification changes.

Funding

This research has not been financed.

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.

Data availability

The research base data cannot yet be shared as it is in use in other ongoing research.

References

[1] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature. 521 (2015) 436–444, https://doi.org/10.1038/nature14539.

[2] R. Mu, X. Zeng, A review of deep learning research, KSII Trans. Internet Inf. Syst. 13 (2019) 1738–1764, https://doi.org/10.3837/tiis.2019.04.001.

[3] S. Sengupta, S. Basak, P. Saikia, S. Paul, V. Tsalavoutis, F. Atiah, V. Ravi, A. Peters, A review of deep learning with special emphasis on architectures, applications and recent trends, Knowl.-Based Syst. 194 (2020), 105596, https://doi.org/10.1016/j.knosys.2020.105596.

[4] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, Y. Miao, Review of image classification algorithms based on convolutional neural networks, Remote Sens. 13 (2021) 4712, https://doi.org/10.3390/RS13224712.

[5] R. Girshick, J. Donahue, T. Darrell, J. Malik, Region-based convolutional networks for accurate object detection and segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2016) 142–158, https://doi.org/10.1109/TPAMI.2015.2437384.

[6] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431–3440, https://doi.org/10.1109/CVPR.2015.7298965.

[7] D. Ciregan, U. Meier, J. Schmidhuber, Multi-column deep neural networks for image classification, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012, pp. 3642–3649, https://doi.org/10.1109/CVPR.2012.6248110.

[8] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778, https://doi.org/10.1109/CVPR.2016.90.

[9] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, ArXiv Preprint ArXiv:1409.1556, 2015, pp. 1–14, https://doi.org/10.48550/arXiv.1409.1556.

[10] C. Szegedy, A. Toshev, D. Erhan, Deep neural networks for object detection, Adv. Neural Inf. Proces. Syst. 26 (2013) 2553–2561, in: https://proceedings.neurips.cc/paper/2013/file/f7cade80b7cc92b991cf4d2806d6bd78-Paper.pdf (accessed July 27, 2021).

[11] D. Shen, G. Wu, H.-I. Suk, Deep learning in medical image analysis, Annu. Rev. Biomed. Eng. 19 (2017) 221–248, https://doi.org/10.1146/annurev-bioeng-071516-044442.

[12] M. Bakator, D. Radosav, Deep learning and medical diagnosis: a review of literature, Multimodal Technol. Interact. 2 (2018) 47, https://doi.org/10.3390/mti2030047.

[13] L. Cai, J. Gao, D. Zhao, A review of the application of deep learning in medical image classification and segmentation, Ann. Transl. Med. 8 (2020) 713, https://doi.org/10.21037/ATM.2020.02.44.

[14] I. Sa, Z. Ge, F. Dayoub, B. Upcroft, T. Perez, C. McCool, Deepfruits: a fruit detection system using deep neural networks, Sensors. 16 (2016) 1222, https://doi.org/10.3390/S16081222.

[15] E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez, Convolutional neural networks for large-scale remote-sensing image classification, IEEE Trans. Geosci. Remote Sens. 55 (2017) 645–657, https://doi.org/10.1109/TGRS.2016.2612821.

[16] I. Gil Leiva, P. Díaz Ortuño, J.V. Rodríguez Muñoz, Técnicas y usos en la clasificación automática de imágenes/Techniques and uses in the automatic classification of images, in: J. Tramullas, P. Garrido-Picazo, G. Marco-Cuenca (Eds.), Actas Del IV Congreso ISKO España y Portugal 2019, 2020, pp. 11–26, https://doi.org/10.5281/zenodo.3733409.

[17] T.D. Akinosho, L.O. Oyedele, M. Bilal, A.O. Ajayi, M.D. Delgado, O.O. Akinade, A. A. Ahmed, Deep learning in the construction industry: a review of present status and future innovations, J. Build. Eng. 32 (2020), 101827, https://doi.org/10.1016/j.jobe.2020.101827.

[18] A.S. Greeshma, J.B. Edayadiyil, Automated progress monitoring of construction projects using machine learning and image processing approach, Mater. Today: Proc. 65 (2022) 554–563, https://doi.org/10.1016/J.MATPR.2022.03.137.

[19] J. Llamas, P.M. Lerones, R. Medina, E. Zalama, J. Gómez-García-Bermejo, Classification of architectural heritage images using deep learning techniques, Appl. Sci. 7 (2017) 992, https://doi.org/10.3390/APP7100992.

[20] M.H. Abed, M. Al-Asfoor, Z. Hussain, Architectural heritage images classification using deep learning with CNN, in: 2nd International Workshop on Visual Pattern Extraction and Recognition for Cultural Heritage. 2602, 2020, pp. 1–12, https://doi.org/10.13140/RG.2.2.20042.11204.

[21] M.J. Carretero-Ayuso, A. Moreno-Cansado, Análisis estadístico nacional sobre patologías en edificación: 2008 a 2010/National statistical analysis on pathologies in buildings: 2008 to 2010 (Technical report), Mussat Fundation Report, 2013, pp. 1–27. https://fundacionmusaat.musaat.es/media/pdf/publicaciones/Resumen_patologias_I_1.pdf (accessed February 10, 2021).

[22] W.R.L. da Silva, D.S. de Lucena, Concrete cracks detection based on deep learning image classification, Proceedings. 2 (2018) 489, https://doi.org/10.3390/ICEM18-05387.

[23] H. Perez, J.H.M. Tah, A. Mosavi, Deep learning for detecting building defects using convolutional neural networks, Sensors. 19 (2019) 3556, https://doi.org/10.3390/s19163556.

[24] 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.

[25] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (1943) 115–133, https://doi.org/10.1007/BF02478259.

[26] F. Attneave, M. B, Review of the organization of behavior; A neuropsychological theory, by D. O. Hebb, Am. J. Psychol. 63 (1950) 633–642, https://doi.org/10.2307/1418888.

[27] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, Psychol. Rev. 65 (1958) 386–408, https://doi.org/10.1037/h0042519.

[28] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature. 323 (1986) 533–536, https://doi.org/10.1038/323533a0.

[29] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (1998) 2278–2323, https://doi.org/10.1109/5.726791.

[30] P.P. Shinde, S. Shah, A review of machine learning and deep learning applications, in: 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1–6, https://doi.org/10.1109/ICCUBEA.2018.8697857.

[31] R. Zhang, W. Li, T. Mo, Review of deep learning, ArXiv Preprint ArXiv:1804.01653, 2018, pp. 385–397, https://doi.org/10.48550/arXiv.1804.01653.

[32] N. Aloysius, M. Geetha, A review on deep convolutional neural networks, in: 2017 International Conference on Communication and Signal Processing (ICCSP), 2017, pp. 0588–0592, https://doi.org/10.1109/ICCSP.2017.8286426.

[33] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, L. Fei-Fei, ImageNet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (2015) 211–252, https://doi.org/10.1007/s11263-015-0816-y.

[34] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980–2988, https://doi.org/10.1109/ICCV.2017.322.

[35] Z. Zou, Z. Shi, Y. Guo, J. Ye, S. Member, Object Detection in 20 years: A Survey, ArXiv Preprint ArXiv:1905.05055, 2019, pp. 1–39, https://doi.org/10.48550/ARXIV.1905.05055.

[36] K. He, X. Zhang, S. Ren, J. Sun, Spatial pyramid pooling in deep convolutional networks for visual recognition, IEEE Trans. Pattern Anal. Mach. Intell. 37 (2015) 1904–1916, https://doi.org/10.1109/TPAMI.2015.2389824.

[37] R. Girshick, Fast R-CNN, 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1440–1448, https://doi.org/10.1109/ICCV.2015.169.

[38] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, in: Advances in Neural Information Processing Systems, 2015, pp. 1–14, https://doi.org/10.48550/arXiv.1506.01497. ArXiv: 1506.01497.

[39] J. Dai, Y. Li, K. He, J. Sun, R-FCN: Object detection via region-based fully convolutional networks, in: Advances in Neural Information Processing Systems, 2016, pp. 1–11, https://doi.org/10.48550/arXiv.1605.06409. ArXiv:1605.06409.

[40] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788, https://doi.org/10.1109/CVPR.2016.91.

[41] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, A.C. Berg, SSD: single shot multibox detector, computer vision - ECCV, Lect. Notes Comput. Sci 9905 (2016) 21–37, https://doi.org/10.1007/978-3-319-46448-0_2.

[42] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, Focal loss for dense object detection, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980–3007, https://doi.org/10.1109/ICCV.2017.324.

[43] P. Jiang, D. Ergu, F. Liu, Y. Cai, B. Ma, A review of yolo algorithm developments, Procedia Comput. Sci. 199 (2022) 1066–1073, https://doi.org/10.1016/J.PROCS.2022.01.135.

[44] D. Thuan, Evolution of Yolo Algorithm and Yolov5: The State of the Art Object Detection Algorithm, Oulu University of Applied Sciences, 2021, pp. 1–61. https://www.theseus.fi/bitstream/handle/10024/452552/Do_Thuan.pdf?sequence=2 (accessed January 27, 2022).

[45] M.J. Carretero-Ayuso, M.P. Sáez-Pérez, Construction flaws in facing brick facades and the risk of associated litigation, J. Build. Eng. 33 (2021), 101633, https://doi.org/10.1016/J.JOBE.2020.101633.

[46] M.J. Carretero-Ayuso, C.E. Rodríguez-Jiménez, D. Bienvenido-Huertas, J. J. Moyano, Interrelations between the types of damages and their original causes in the envelope of buildings, J. Build. Eng. 39 (2021), 102235, https://doi.org/10.1016/J.JOBE.2021.102235.

[47] J.M. Rincón, M. Romero, Basis and classification of efflorescences in construction bricks, Mater. Constr. 50 (2000) 63–69, https://doi.org/10.3989/mc.2000.v50.i260.391.

[48] F.E. Salvador Esteve, Estudio de lesiones en fachadas de ladrillo cara vista. Análisis de casos en Yecla/Study of lesions in exposed brick facades, Anal. Cases Yecla (2015) 1–279. http://hdl.handle.net/10045/48917 (accessed July 27, 2021).

[49] A. García Verduch, V. Sanz Solana, in: F.E. Iberica (Ed.), Velos, florescencias y manchas en obras de ladrillo, 1999, pp. 1–301. Castellón, Spain. ISBN:84–87683-10–X, https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=Sanz%2C+V.%2C+%26+García%2C+A.+%281999%29.+Velos%2C+florescencias+y+ma

nchas+en+obras+de+ladrillo.+Faenza+Editrice+Iberica%2C+Castellón+%28España%29.&btnG= (accessed February 18, 2021).
[50] J.M. Rincón, M. Romero, Prevención y eliminación de eflorescencias en la restauración de ladrillos de construcción/ Prevention and elimination of efflorescence in the restoration of building bricks, Mater. Constr. 51 (2001) 73–78, https://doi.org/10.3989/mc.2001.v51.i261.382.
[51] J. Monjo Carrió, in: E. Munilla-Lería (Ed.), Patología de cerramientos y acabados arquitectónicos/Pathology of architectural closures and finishes, 1994, pp. 1–399. Madrid. ISBN; 84-89150-02-8, https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=J.M.+Carrió%2C+Patología+de+cerramientos+y+acabados+arquitectónicos&btnG (accessed February 18, 2021).
[52] C. Broto, A. Mostaedi, Enciclopedia Broto de patologías de la construcción: Elementos constructivos I: cerramientos exteriores, interiores, puertas, ventanas, y cristales/Broto Encyclopedia of construction pathologies: Construction elements I: exterior and interior enclosures, in: Link International (Ed.), Barcelona, Spain, 2006, pp. 1–311. ISBN: 84-96424-38–3, https://scholar.google.com/scholar_lookup?title=EnciclopediaBrotodelaspatologíasdelaconstrucción&publication_year=2006&author=C.Broto (accessed February 18, 2021).
[53] Government of Spain, Order of July 27, 1988 Approving the General Specifications for the recepTion of Ceramic Bricks in Construction Works RL-88, BOE 03/09/1988, n.185, 1988, pp. 23921–23924. https://www.boe.es/eli/es/o/1988/07/27/(4) (accessed September 28, 2021).
[54] Government of Spain, Royal Decree 1371/2007, of October 19, which Approves the Basic Document "DB-HR Protection against Noise" of the Technical Building Code and Modifies Royal Decree 314/2006, of March 17, by which Approves the Technical Code, BOE 23/10/2007.n. 254, 2007, pp. 42992–43045. https://www.boe.es/eli/es/rd/2007/10/19/1371 (accessed September 28, 2021).
[55] G. Jocher, A. Stoken, J. Borovec, NanoCode012, A. Chaurasia, TaoXie, L. Changyu, A. V, Laughing, tkianai, yxNONG, A. Hogan, lorenzomammana and, AlexWang1900 and, J. Hajek, L. Diaconu, Marc, Y. Kwon, oleg, Wanghaoyang0106, Y. Defretin, A. Lohia, Ml5ah, B. Milanko, B. Fineran, D. Khromov, D. Yiwei, Doug, D. And, F. Ingham, Ultralytics/yolov5: v5.0 - YOLOv5-P6 1280 Models, AWS, Supervise.Ly and YouTube Integrations (v5.0), Zenodo, 2021, pp. 1–5, https://doi.org/10.5281/zenodo.4679653.