Crack detection in buildings using the YOLO v8 network

W. S. Ribeiro[1]* , J. Zanetti[1] , L. B. Totola[1] , S. A. C. Junqueira[1] , P. H. P. Lauff[1]
*Contact author: weiglasribeiro@gmail.com
DOI: https://doi.org/10.21041/ra.v14i3.76

ABSTRACT

The objective of this study is to develop and apply deep neural networks for the automation of crack detection in buildings. The methodology involved training the YOLO v8 network with images collected from the internet, aiming to identify and locate cracks in real time. The model obtained 80% accuracy in validation with images not used in training, despite performance limitations in Google Collab. These limitations included restrictions on the execution environment, and the model is specific to cracks. The originality of the tool lies in its relevance for the automated detection of cracks, with the potential to extend to other pathological manifestations. It is concluded that the application of deep neural networks offers an efficient solution for the identification of problems in buildings.

Keywords: pathological manifestations; building construction; crack detection; image analysis; YOLO v8.

---

[1] Department of Engineering, Multivix Vila Velha College, Vila Velha, Brazil.

Contribution of each author

In this work, the authors Weiglas Soriano Ribeiro and Juliette Zanetti participated in the data processing in 70% and 30%, respectively. Weiglas Soriano Ribeiro, Juliette Zanetti and Lucas Broseghini Totola contributed to the writing of the scientific article in the following areas: original idea, experimentation, model development, data collection, article writing and discussion of results. The percentages of contribution to the writing of the article were: Weiglas Soriano Ribeiro with 20%, Juliette Zanetti with 40% and Lucas Broseghini Totola with 40%. Sérgio Ândrigo Colaço Junqueira and Pedro Henrique Pina Lauff contributed equally (50% each) to the acquisition of the database and carrying out the tests.

Discussions and subsequent corrections to the publication

Any dispute, including the replies of the authors, will be published in the second issue of 2025 provided that the information is received before the closing of the first issue of 2025.

Detecção de trincas em edificações utilizando a rede YOLO v8

RESUMO

O objetivo deste estudo é desenvolver e aplicar redes neurais profundas para a automação da detecção de trincas em edificações. A metodologia envolveu o treinamento da rede YOLO v8 com imagens coletadas da internet, visando identificar e localizar trincas em tempo real. O modelo obteve 80% de acurácia na validação com imagens não utilizadas no treinamento, apesar das limitações de desempenho no Google Collab. Essas limitações incluíram restrições no ambiente de execução, e o modelo é específico para trincas. A originalidade da ferramenta está na sua relevância para a detecção automatizada de trincas, com potencial de se estender a outras manifestações patológicas. Conclui-se que a aplicação de redes neurais profundas oferece uma solução eficiente para a identificação de problemas em edificações.
Palavras-chave: manifestações patológicas; construção civil; detecção de rachaduras; análise de imagem; YOLO v8.

Detección de grietas en edificaciones mediante la red YOLO v8

RESUMEN

El objetivo de este estudio es desarrollar y aplicar redes neuronales profundas para automatizar la detección de grietas en edificios. La metodología consistió en entrenar la red YOLO v8 con imágenes recopiladas de Internet, con el objetivo de identificar y localizar grietas en tiempo real. El modelo logró una precisión de validación del 80 % con imágenes no utilizadas en el entrenamiento, a pesar de las limitaciones de rendimiento en Google Collab. Estas limitaciones incluían restricciones sobre el entorno de ejecución y que el modelo fuera específico para las grietas. La originalidad de la herramienta radica en su relevancia para la detección automatizada de fisuras, con potencial de extensión a otras manifestaciones patológicas. Se concluye que la aplicación de redes neuronales profundas ofrece una solución eficiente para la identificación de problemas en las edificaciones.
Palabras clave: manifestaciones patológicas; construcción de edificios; detección de grietas; análisis de imágenes; YOLO v8.

Ribeiro, W. S., Zanetti, J., Totola, L. B., C. J. unqueira, S. A., Lauff, P. H. P.

1. INTRODUCTION

In civil engineering, structural pathology focuses on understanding the causes and mechanisms of structural degradation that can occur during the design, construction, and use of a building (Caporrino, 2018). Pathological manifestations signal a failure in a structure's performance, potentially compromising its durability, safety, and functionality. Therefore, accurately identifying and diagnosing the origins and development of these anomalies is essential for implementing effective corrective measures. (Bolina et al., 2018).

Cracks are one of the most common forms of pathological manifestation in buildings and can indicate structural risks. The causes of cracks are diverse, including the inherent nature of the material, such as reinforced concrete, which has low tensile strength and is prone to shrinkage, as well as deficiencies in the design and construction phases, differential settlements, the use of low-quality materials, and processes inherent to the aging of the structure itself (De Souza and Ripper, 1998). Furthermore, cracks constitute a visual aspect that leads to property devaluation, creates insecurity among users, and serves as a warning sign, acting as facilitators for weathering effects on structures.

Kung et al. (2018) and Yu (2022) report that conventional approaches for detecting pathological manifestations are carried out through manual inspections and photographic records of buildings. When the building consists of a multi-story structure, or in cases of special works such as bridges and walkways, the data collection for inspections becomes costly and labor-intensive, potentially even posing safety risks (Kung et al., 2021; Ribeiro et al., 2020). In this context, the use of automated artificial intelligence approaches associated with unmanned aerial vehicles (UAVs) for the collection and interpretation of images is becoming an important tool in the identification of pathological manifestations. UAVs can capture photographs and serve as tools for accessing and operating other investigative technologies, such as thermography. According to Kneipp (2018), the advantages of using UAVs for inspection include time optimization, increased operator comfort, the ability to investigate confined spaces, and the capability to reach great heights without the need for workers to be tied to ropes or require scaffolding.

Cha et al. (2017) emphasize that crack detection through visual inspection can be a complex procedure, influenced by the number of cracks and access difficulties, and is considerably affected to observer subjectivity. Given these challenges, various methods have been proposed to automate this process using advanced image processing techniques. However, the effective application of these techniques faces obstacles in adverse conditions, such as fluctuations in lighting and variations in material textures.

Some neural networks, such as YOLO (You Only Look Once), are applied in image classification, predicting objects in an image and highlighting them with bounding boxes. Introduced by Redmon et al. (2015), YOLO is a real-time object recognition system known for its superior accuracy and speed compared to other recognition systems.

In this context, the present study aims at the automated detection of cracks in buildings using machine learning, through the development of artificial neural networks using the YOLO model. By collecting images and training the neural networks, this study aims to analyze the potential and confirm the feasibility of this tool for automating this process in civil construction.

2. THEORETICAL REFERENCE

2.1 YOLO-v8

Computer vision systems have emerged remarkably in the contemporary scenario, playing an important role in vehicle automation, industrial robotization and hospital devices. One of its prominent applications is the ability to perform diagnoses through image exams, representing significant advances in machine automation and in resolving various challenges (Barelli, 2018).

Ribeiro, W. S., Zanetti, J., Totola, L. B., C. Junqueira, S. A., Lauff, P. H. P.

As highlighted by Mantripragada (2020), object detection technology encompasses two fundamental tasks: identifying the class and determining the location of objects. This technology demonstrates wide applicability and can be used in both static images and real-time videos.

Redmon et al. (2016) highlight the application of the YOLO algorithm, whose central objective is the classification and detection of objects. This algorithm allows obtaining the position and category assigned to the object identified in the image in which the prediction was made. Using a single convolutional neural network (CNN), YOLO simultaneously anticipates multiple bounding boxes and the classification probabilities associated with those boxes.

To understand how the YOLO algorithm works, it is essential to define what is being predicted: the class of an object and the bounding box that specifies its location. Each bounding box is characterized by four elements, as highlighted by Swiezewski (2020):
- Center of the bounding box (bx, by);
- Width (PB);
- Height (bh); and
- The value c corresponds to a class of an object (such as: car, traffic light, cracks, etc.).

In addition to these elements, there is a need to predict the value of p_c (class probability): a measure that estimates the probability of there being an object contained in the bounding box. Figure 01 exemplifies object detection in the Yolo algorithm.
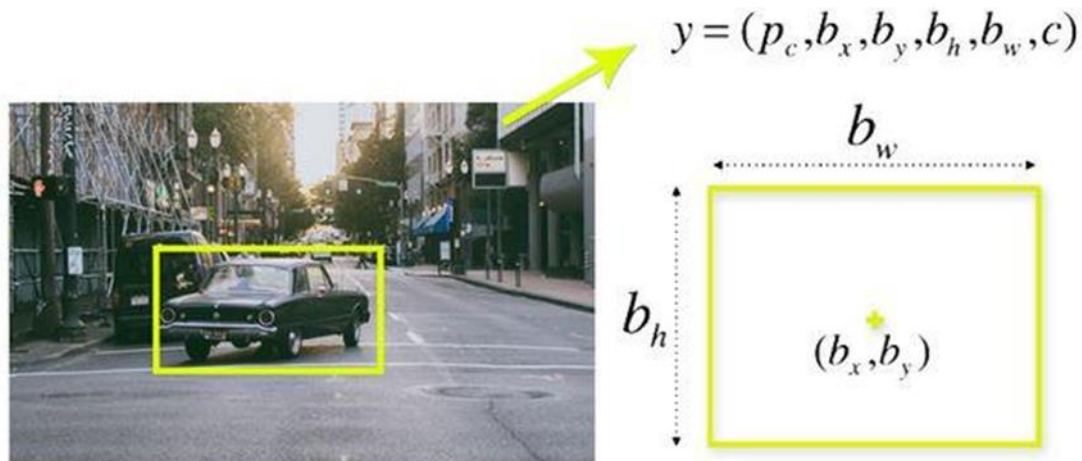

Figure 1. Exemplifying object detection in the YOLO algorithm.
Source: Swiezewski (2020).

In the YOLO algorithm, the image is divided into cells (Figure 2), each responsible for predicting up to 5 bounding boxes for objects. However, many of these cells and boxes do not contain an object. To deal with this, p_c values are used to remove boxes with a low chance of containing an object and boxes with large overlap, through a process called non-maximum suppression. During detection, multiple bounding boxes can be detected for each class. To reduce the number of detected boxes and remove overlaps, the non-maximum suppression (NMS) algorithm is applied, as shown in Figure 1. NMS compares the properties of each box, such as the confidence score, and keeps only the most confident one (Redmon et al., 2016).
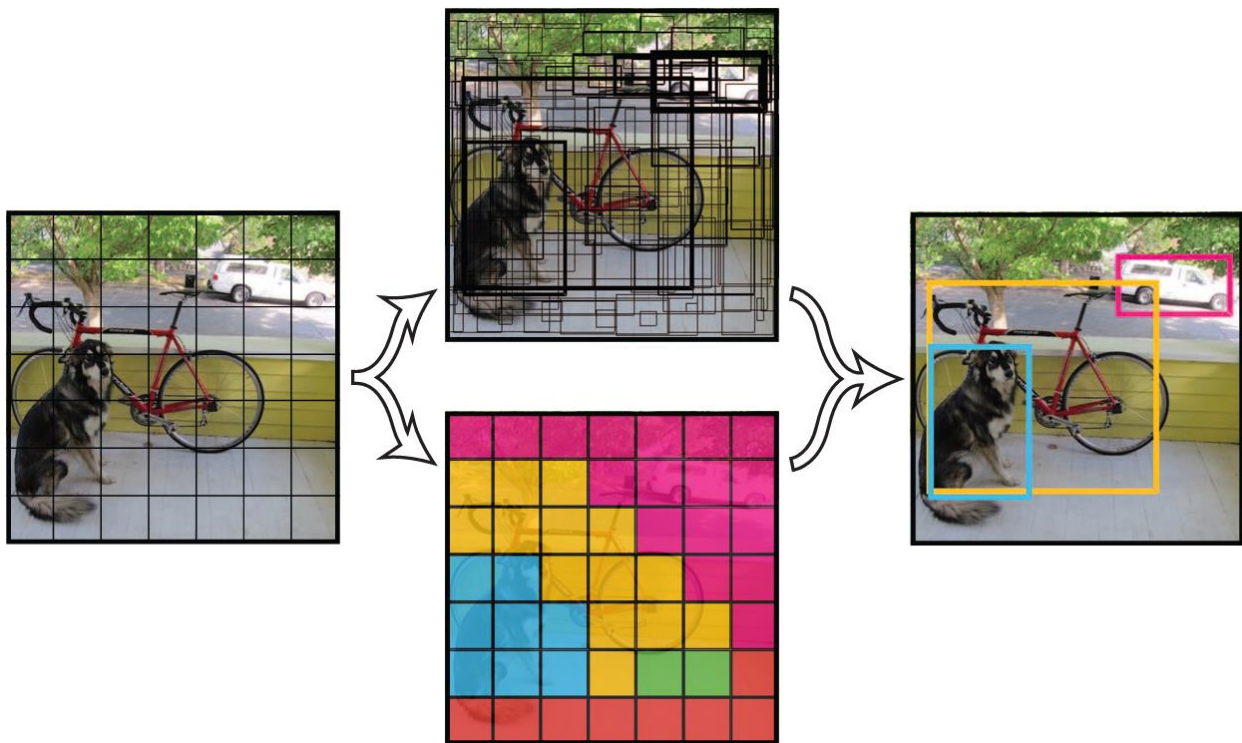
Figure 2. NMS algorithm in action after detecting several bounding boxes.
Source: Bavaresco (2023).

The YOLO method, proposed by Redmon et al. (2016), represents a significant reformulation in object detection, transforming it into a regression problem. This unique approach starts exclusively from the pixels of an image, resulting in predictions that cover the probabilities per class, the coordinates and the dimensions that delimit the objects in question. YOLO's simplicity is remarkable as it takes an end-to-end approach through a single CNN. In addition to this simplicity, it stands out for presenting competitive performance in terms of time efficiency.

Since its creation by Redmon et al. (2016), until the last revision in 2024, the YOLO algorithm went through nine iterations, with YOLO-v8 being used in the present study. Throughout this period, the model's architecture was continually improved to ensure efficiency, better performance and superiority over previous versions (Hussain, 2023).

Launched by Ultralytics in January 2023, YOLO-v8 stands out for offering optimized performance in terms of speed and accuracy. This release supports multiple artificial intelligence (AI) vision tasks, covering tracking, classification, pose segmentation, detection, and segmentation. YOLO-v8's remarkable flexibility allows its users to leverage its features across different hardware platforms (Batistoti, 2023).

Given the growing need for effective automated techniques for mapping pathological manifestations in the civil engineering field, a few studies have employed CNNs. Ekanayake (2022) developed a deep learning-based YOLO algorithm that provides an automated monitoring tool to ensure the sustainability of buildings. Kung et al. (2021) and Woo et al. utilized unmanned aerial vehicles (UAVs) to detect defects in buildings and developed CNNs for crack detection.

Next, the methodology applied in the present study is presented in detail, highlighting the relevance of YOLO-v8 in this specific context.

3. METHODOLOGY

To facilitate understanding of the methodology adopted, a flowchart of the activities carried out is presented in Figure 3. Within the scope of this study, images of geometric cracks in reinforced concrete structures, masonry walls, floors and concrete walls were selected, totaling 303 samples to compose the database. The causes of cracks were not the subject of study in this research. Using the Roboflow software, each crack present in the images was manually delimited, resulting in the constitution of the set of samples. It is noteworthy that, during this phase, the Image augmentation feature was applied, which introduces random variations to the original images, generating new training instances with characteristics different from the original ones. The implementation of Ultralytics and the download of the YOLO v8 project were carried out in the Google Collab environment. Subsequently, training began, comprising a cycle of 400 epochs.
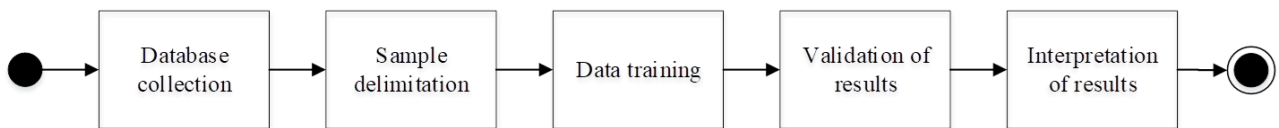
Figure 3. Flowchart of the process steps.
Source: The authors.

The available database was divided into two sets: a training set for adjusting the model parameters (75% of the total sample), and a validation set of images (25% of the total sample) to test the robustness of the proposed models. After the training stage, a test was carried out to validate the results, using images not included in the training base. Finally, the results obtained were compiled and interpreted.

In the present work, the mean average precision (mAP) values were analyzed, a metric used in object detection that indicates the evolution of training. After several iterations, the accuracy is represented by a graph that seeks to approach 100%. As training progresses, the aim is to reduce the sizes of the bounding boxes to obtain better average precision, thus increasing the value of the network's accuracy, which reflects the increasing performance of the network (Divvala, 2015).

4. RESULTS AND DISCUSSION

This topic presents training statistics, as well as the results of processing some images that do not belong to the training data set.

Figure 4 shows the result of processing a crack image by the YOLO network, trained with 80% accuracy. This value indicates that the network makes few false marking errors and, at the same time, does not fail to mark the necessary objects (cracks). This image was strategically selected, containing only one crack in the area, to analyze the behavior in the simplest way possible. The network presents a significantly precise crack delimitation. Notably, minor plaster imperfections resembling cracks, as seen in Figure 4, were not detected by the network, as they are not considered pathological manifestations.

Figure 4. Result of processing an image with 80% accuracy.
Source: The authors.

Figure 5 shows a more complex case, with more than one crack. The YOLO network detected two cracks with accuracies greater than 42% and 79%, respectively. It is observed that, given the distribution of cracks in the image provided, the 42% result returned statistics below expectations, indicating the need for improvements in training. However, even with relatively low precision, YOLO was able to correctly identify the two cracks in the image.
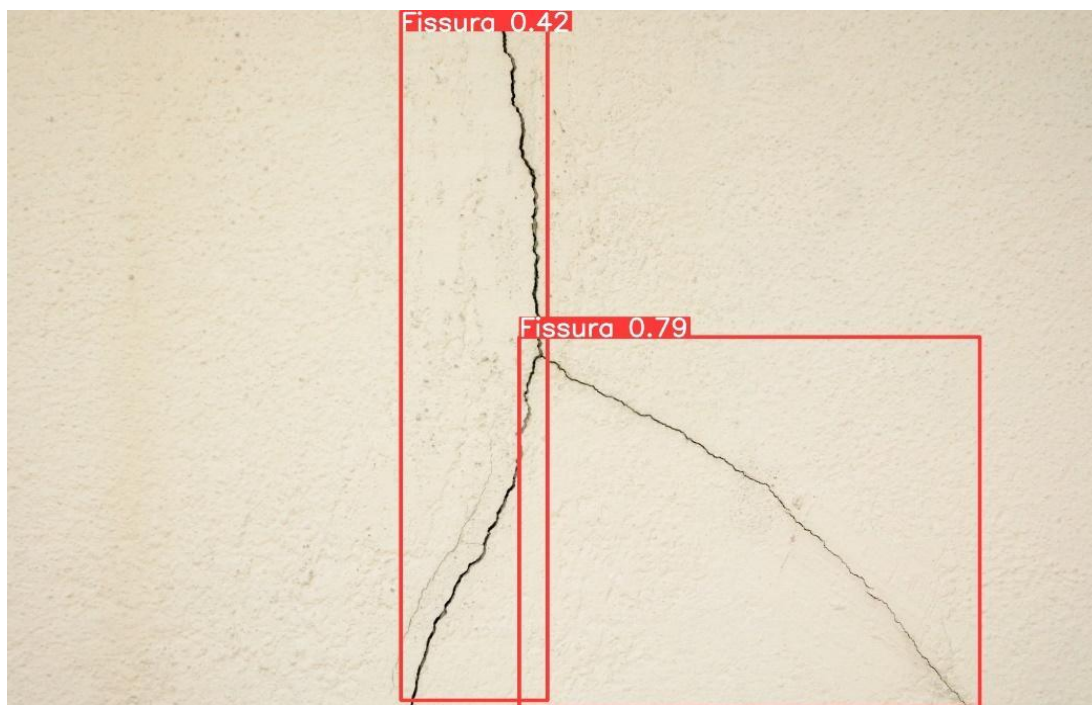
Figure 5. Result of processing an image with 42% and 79% accuracy.
Source: The authors.

Ribeiro, W. S., Zanetti, J., Totola, L. B., C. Junqueira, S. A., Lauff, P. H. P.

In Figure 6, the YOLO network was applied to a masonry wall with variations in color tone. The results show that the network detected two cracks with an accuracy of 58%. However, around the colored wall, no cracks were identified, indicating the influence of lighting and shadows on the results, as noted by Cha and Choi (2017). Therefore, a diverse and extensive database is needed to improve training across various scenarios. Finally, it is interesting to note that the accuracy was the same in the two identified boxes, suggesting a consistent pattern of behavior, as both are characterized in the same direction.

Figure 6. Result of processing an image with 58% accuracy.
Source: The authors.

The evaluation scores for YOLOv8 are presented in Figure 7, where (a) to (e) refer to the training phase and (f) to (j) refer to the validation phase. The loss observed in Figure 7 (a) is related to the bounding boxes in relation to the objects found by the algorithm, presenting a loss associated with the central coordinates of the object and the ends of the boxes. Figure 7 (b) shows the loss associated with the classification of boxes in relation to the objects found, referring to the IoU. Finally, Figure 7 (c) presents the loss associated with Local Density-Free, whose function is to adjust the trained model and regulate the density of objects in different regions of the bounding boxes, especially when objects are close to each other.
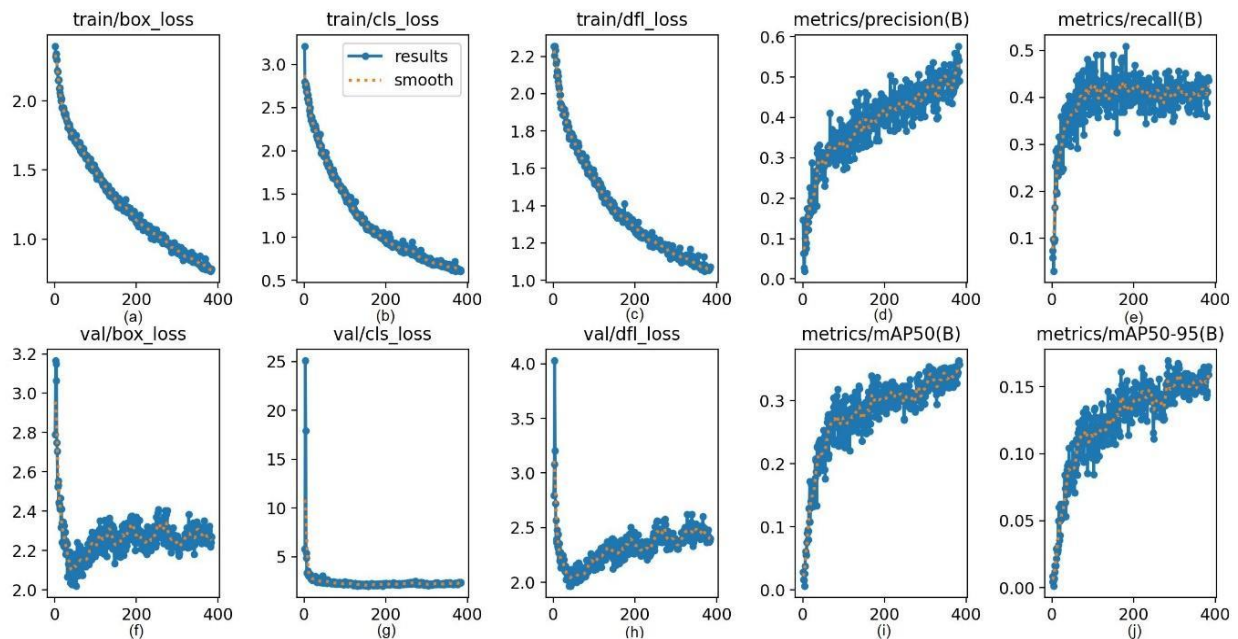
Figure 7. Results of the recall precision, mAP 50% and mAP 50-95% performance metrics for the YOLOv8 algorithm. Source: The authors.

Thus, Figures 7 (a), (b) and (c) show an inversely proportional relationship between the number of epochs trained and losses, indicating that the network training performance improves over the 400 epochs. Figures 7(d) and (e) refer to the precision and recall metrics, respectively. It is noted that, with the increase in the number of epochs, the values of the performance metrics also increase, presenting a directly proportional relationship.

Figures 7 (f) to (j) follow the same reasoning as the training phase, but in the validation phase. It is also possible to observe a good performance of the classifier at this stage, although the mAP50 and mAP50-95 metrics show fluctuations over the periods.

This study aimed to demonstrate the application of the YOLO v8 neural network for crack detection in buildings, revealing the effectiveness and predictability of this technology to automate the inspection process. The results obtained indicate that YOLO v8 is a promising tool for this task, offering an automated solution that can increase the efficiency and accuracy of pathology assessments in buildings.

The quality of the results presented is directly related to the quality and quantity of data used to train the network. Images with overlapping cracks, for example, resulted in lower accuracies, showing that the presence of multiple overlapping cracks can confuse the algorithm and reduce its detection capacity. To mitigate these limitations, it is essential to expand the image dataset, including a wider range of cases with overlaps and variations. A more robust and diverse database will allow the YOLO v8 network to learn to distinguish between different types of pathological manifestations, improving detection accuracy.

Furthermore, using a dedicated machine for training, instead of a free environment like Google Colab, would allow for more efficient processing with a higher number of epochs. This increase in the number of epochs can lead to a significant improvement in the accuracy of the results, providing a more refined and reliable model.

The results of this study indicate that the YOLO v8 algorithm is reliable under conditions similar to those used in the tests. However, to achieve even more accurate detection, a research objective that seeks mAP (mean average accuracy) values above 90% is necessary. Higher mAP values provide greater reliability and robustness in the detections made by the network, allowing for more effective application in real-world scenarios.

Ribeiro, W. S., Zanetti, J., Totola, L. B., C. Junqueira, S. A., Lauff, P. H. P.

The need to identify early aging of pathological manifestations highlights the importance of automated and continuous monitoring tools. The use of neural networks like YOLO v8 can transform the inspection process, making it less dependent on manual assessments, which are often slow and costly. Automation not only reduces the cost and time required to detect cracks and other pathologies, but also increases the frequency and accuracy of inspections, contributing to the maintenance and safety of buildings.

Therefore, the application of YOLO v8 in the building pathology industry demonstrates a significant advancement in the way inspections are performed, highlighting the convenience and utility of automated tools for continuous monitoring of structural integrity. Continuous development and improvement of detection algorithms is essential to achieve a level of accuracy that allows a complete and reliable assessment of building conditions.

5. CONCLUSIONS

Neural networks play an essential role in pattern recognition and anomaly localization. This study demonstrated that the YOLO v8 network is a highly effective tool for automated crack detection in buildings. The accuracy of the results is closely linked to the quality and diversity of the training data. Overlapping images of cracks can compromise accuracy, but this limitation can be overcome with a more robust dataset and an improved training environment. The algorithm proved to be reliable under the tested conditions and has great potential for adaptation to detect other structural pathologies.

For future work, it is recommended to increase the number of training images to improve the accuracy of the model. Furthermore, it is proposed to investigate the application of the YOLO network for real-time crack detection during field data acquisition.

6. ACKNOWLEDGMENTS

7. REFERENCES

Barelli, F. (2018), "*Introduction to Computer Vision: A practical approach with Python and OpenCV*". Code House.

Batistóti, J.O. (2023), "*Remote sensing in the identification and characterization of crops of zootechnical interest*". Thesis (PhD) – Faculty of Veterinary Medicine and Animal Science, Federal University of Mato Grosso do Sul, Campo Grande – MS.

Bavaresco, L. (2023), "*Instance segmentation for estimating fish length using artificial intelligence techniques*". Course completion work (graduation) - Federal University of Santa Maria, Technological Center, Computer Engineering Course, RS.

Bolina, F. L., Tutikian, B. F., Helena, P. (2019). "*Structural pathology*". Text Workshop.

Caporrino, C. F. (2018). *"Pathology in Freemasonry"*. 2nd edition. São Paulo: Oficina de Textos.

Cha, Y.-J., Choi, W., Büyüköztürk, O. (2017). "*Deep learning-based crack damage detection using convolutional neural networks*". Computer Aided Civil and Infrastructure Engineering, 32(5), p. 361–378.

De Souza, V. C. M., Ripper, T. (1998). *"Pathology, recovery and reinforcement of concrete structures"*. Pini.

Ribeiro, W. S., Zanetti, J., Totola, L. B., C. J. unqueira, S. A., Lauff, P. H. P.

Divvala, S., Redmon, J., Girshick, R., Farhadi, A. (2015). *"You only look once: unified real-time object detection"*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Ekanayake, B. (2022). *"A deep learning-based construction defect detection tool for sustainability monitoring"*. In: 10th World Construction Symposium.

Hussain, M. (2023). *"YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature towards digital manufacturing and industrial defect detection"*. Machines, vol. 11, no. 7, 2023. https://doi.org/10.3390/machines11070677

Kneipp, R. B. (2018). *"The state of the art in the use of Drones for Naval and Offshore Inspection"*. 81f. Dissertation - Federal University of Rio de Janeiro, Rio de Janeiro.

Kung, R.-Y., Pan, N.-H., Wang, C. C. N., Lee, P.-C. (2021). *"Application of Deep Learning and Unmanned Aerial Vehicles in Building Maintenance"*. Advances in Civil Engineering, Volume 2021, Issue 1, 5598690. https://doi.org/10.1155/2021/5598690

Mantripragada, M. (2020). *"Digging deeper into YOLO V3 – A practical guide Part 1"*. Available at: https://towardsdatascience.com/digging-deep-into-yolo-v3-a-hands-on-guide-part-1 -78681f2c7e29

Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). *"You Only Look Once: Unified Real-Time Object Detection"*. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), https://doi.org/10.1109/CVPR.2016.91

Ribeiro, D., Santos, R., Shibasaki, A., Montenegro, P., Carvalho, H., Calçada, R. (2020), *Remote inspection of RC structures using unmanned aerial vehicles and heuristic image processing*, Engineering Failure Analysis, Volume 117, 104813, ISSN 1350-6307, https://doi.org/10.1016/j.engfailanal.2020.104813

Swiezewski, J. (2020). *"Yolo Algorithm and Yolo Object Detection: An Introduction"*. Available at: <https://appsilon.com/object-detection-yolo-algorithm>.

Woo, H. J., Seo, D. M., Kim, M. S., Park, M. S., Hong, W. H., Baek, S. C. (2022). *"Localization of cracks in concrete structures using an unmanned aerial vehicle"*. Sensors, 22(17), 6711, https://doi.org/10.3390/s22176711

Yu, Z. (2022). *"Deep learning approach based on YOLO V5s for crack detection in concrete"*. In SHS Web of Conferences (Vol. 144, p. 03015). EDP Sciences.