Available online at www.sciencedirect.com

ScienceDirect

Procedia Structural Integrity 44 (2023) 2028–2035

Structural Integrity Procedia

www.elsevier.com/locate/procedia

XIX ANIDIS Conference, Seismic Engineering in Italy

Using machine learning approaches to perform defect detection of existing bridges

Sergio Ruggieri[a], Angelo Cardellicchio[b], Andrea Nettis[a], Vito Renò[b], Giuseppina Uva[a]*

[a]Department DICATECh, Polytechnic University of Bari, Via Orabona, 4 – 70126, Italy

[b] Institute for Intelligent Industrial Technologies and Systems for Advanced Manufacturing (STIIMA), National Research Council of Italy, Via Amendola, 122 D/O, Bari, Italy

* Correspondence: sergio.ruggieri@poliba.it

Abstract

The paper presents a study about defect detection on structural elements of existing bridges through a machine-learning approach. In detail, the proposed methodology aims to explore the possibility of automatically recognizing defects and damages on bridges' elements, (e.g., cracks, humidity) by employing a training of existing convolutional neural networks on a set of photos. The initial database has been firstly selected and then classified by domain experts according to the requirements of the new Italian Guidelines on structural safety of existing bridges. The results show a good effectiveness and accuracy of the proposed methodology, opening new scenarios for the automatic defect detection on bridges, mainly aimed to support management companies' surveyors in the phase of in-situ structural inspection.

Keywords: Damage Detection; Bridge Inspection; Machine Learning.

Nomenclature

CNN	Convolutional Neural Networks
IoU	Intersection over Union
mAP	map Average Precision
ML	Machine Learning
P	Precision
R	Recall

10.1016/j.prostr.2023.01.259

R-CNN	Regional-based Convolutional Neural Network
RPN	Region Proposal Network
VULMA	VULnerability analysis using MAchine-learning

1. Introduction: new Italian guidelines on structural safety of existing bridges

Several recent events involving the collapse of some strategic bridges in the Italian infrastructural networks have strongly stimulated the public opinion about the safety of existing bridges. In particular, after the collapse of the Polcevera Viaduct (and some minor events such as the Massa Carrara and Torino-Savoia bridges), public institutions and the scientific community have worked to release specific Guidelines on the structural safety of existing bridges (2021). In the new guidelines, a deep focus on the monitoring and maintenance of such structures is made, providing a systematic multilevel procedure to apply to the entire national existing stock. The new Guidelines and the methodology therein will become mandatory for management companies of roads and bridges in the next year, and for this reason, several applications and works are actually under development. According to the new guidelines, the multilevel approach proposes six levels of different complexity to be performed in close sequence. In particular, the first three levels regard a first screening of the infrastructural stock (or part of it), while the second three levels rule the actions to be performed on the infrastructures that have been prioritized after the first three levels. The second three levels are employed to decide if the bridge under investigation should be closed (demolition and reconstruction), retrofitted, or must be subjected to limitations (e.g., traffic, transit of heavy vehicles). Before providing a final response about the future of the bridges, great attention should be devoted to the first three levels. Level 0 aims to define the census of the bridges and viaducts composing the network under study (usually, management companies subdivide the areas according to the regional boundaries). In this phase, many data should be gathered, such as information about the context and the importance within the network, structural features, geometrical and mechanical parameters, constructive details. These data can be found in the original documentation (dating back to the year of construction, if available) or by exploiting existing studies about risks to the environment around the structure. Using the collected information, it is possible to proceed to Level 1, which consists of a survey of the health state of each bridge within the assigned network. This phase is carried out by in-situ inspections in which the surveyors, adequately instructed, observe each element of the bridge (e.g., desk and beams, pillars, supports, abutments) and take note of all the visible structural defects. All observations must be recorded in specific forms (provided within the guidelines), by specifying the extension and the intensity of each defect, to which a specific photo can be associated. Any difference with the original documentation shall be highlighted. In addition, in Level 1, surveyors must identify any possible vulnerability source and boundary condition regarding different risks, such as structural, seismic, hydraulic, geotechnical, and geological. If a specific significant risk affecting the safety of the bridge is identified, the unit must be accurately investigated (additional tests are required) through the procedures in Levels 4, 5, or 6. In the other cases, Level 2 can be performed. Using the information of Level 0 and Level 1, it is possible to define the "risk class", which is a synthetic parameter that accounts for all the analysed risks, each of which is classified with a specific sub-risk class, i.e., structural, seismic, hydraulic, and geological. The Italian Guidelines propose five levels of risk: low, medium-low, medium, medium-high, and high. A scheme is provided with the logical operations to perform for defining the global risk class. Finally, according to the obtained results, the bridge manager will establish the actions to be taken. For example, for a high-risk class, Level 4 shall be carried out (e.g., assessment of transitability, operativity, or retrofit). Also in this case, for the higher risk classes, additional investigations are required (e.g., structural monitoring).

It is worth highlighting that one of the most essential but critical phases of the first three levels are in-situ inspections. Visual observations represent the most direct and fast way to identify issues or potential vulnerabilities in the inspected bridge, but at the same time, the bridges' defect detection based on this approach presents some drawbacks. Among these, a first evident problem is the subjectivity in defining the defect, its extension and severity (although surveyors are usually well-instructed). Moreover, loss of attention can bias the score assigned to a defect, introducing misjudgements due to the human factor. In addition, the inaccessibility of some specific parts of the bridge (e.g., supports) means that some elements cannot be classified. Finally, it is worth remembering that a careful surveyor needs one working day to assess structural elements belonging to about 4/5 bays. For all these reasons, it is highly desirable to have additional tools to support surveyors in the phase of in-situ inspection (this can also optimize management costs) and, the new technology developed in the field of computer vision can perform this arduous task.

In this paper we propose a machine learning (ML) technique for training a reliable and automatic tool for existing bridge defect detection, giving a first contribution to the abovementioned exigency. Starting from a database of photographs, a labelling phase has been carried out on the figures to identify defects in the structural elements of the bridge. Afterward, deep learning-based methods have been used for object detection, exploiting state-of-the-art single-stage detectors, such as YOLOv5. The achieved results show an overall good potentiality for the proposed methodology. Even if both the quantity and the quality of provided data need to be refined and improved, the presented study is the base for a future consolidated tool, which will support surveyors in performing reliable in-situ inspections.

2. State of the art on machine learning and object detection methodologies in structural engineering

During the last years, several applications of ML techniques have been developed by the scientific community in the field of civil and structural engineering. Different ML applications in earthquake engineering, structural properties identification, and structural health monitoring have been developed to propose mathematical tools for solving complex input-output problems (Xie et al., 2020; Sun et al., 2020). Another way to exploit ML applications is to gather information using photographical images. Specifically, images are a principal means for extracting information on structures and infrastructures. For example, Ruggieri et al. (2021) proposed VULMA, a ML tool for defining a simplified vulnerability index of existing buildings starting from a proper dataset (for more information, see Cardellicchio et al., 2022). For the case under study, it is interesting to present a brief overview of the existing work about automatic detection of damages in existing bridges, for which the literature provides a few studies concerning crack detection (e.g., Cha et al., 2017; Xu et al., 2019; Prasanna et al., 2016).

Regarding damage assessment, Potenza et al. (2020) modified the current framework for bridge inspection and condition assessment by introducing a color-based image processing approach to support defect recognition on data collected by unmanned aerial vehicles. Recently, deep learning in damage detection has raised significant research interest. For example, Cha et al. (2018) proposed to use e CNNs for automated detection of five damage types using videos collected during inspection. Zhu et al. (2020) proposed an approach based on transfer learning combined with CNNs for automatically detecting bridge defects.

Concerning the application of CNNs in structural defect detection, CNNs have already been used in literature to perform cracks and damages segmentation. Specifically, a pixel-based analysis via CNNs has been proposed by Zhang et al. (2017) and Yang et al. (2018), while Cha et al. (2017) used a sliding window to perform damage detection on several regions of the original image. Even if these approaches are easy to implement, they cannot deal with scale-variant defects and often imply a high computational cost due to the need to scan the whole image to select damage candidates. Therefore, proper network architectures have been proposed, explicitly tailored for object detection. Currently, two main categories of architectures suited for object detection exist. The first is the two-stage detectors, which rely on a region proposal network (RPN) to extract a series of regions from the overall image as the more likely to contain objects. As for specific application in damage detection, a first example is provided in Girshick et al. (2014), where a Regional-based convolutional neural network (R-CNN) has been used along with morphological post-processing to detect cracked surfaces in concrete bridges. In Cha et al. (2018), the Faster R-CNN technique (Ren et al., 2015) was used to identify five structural surface damages in concrete and steel. A modified version of the Faster R-CNN was also used by Li et al. (2018), which aimed to identify three types of concrete defects. While two-stage detectors usually achieve good accuracy, using an RPN implies an additional computational overhead, severely undermining the performance of the detector. Single-stage detectors, SSD (Liu et al., 2016) and YOLO (Redmon et al., 2016), have been proposed to overcome this issue, considering that RPN is not needed, and then the processing speed is greatly improved and allows real-time detection in several fields. Few studies have already used the SSD technique for damage inspection, such as Maeda et al. (2018), which used it to detect road surface damages. More interest has been directed towards YOLO and its recent versions, e.g., YOLOv2 (Redmon and Fahradi, 2017), YOLOv3 (Redmon and Fahradi, 2018), YOLOv4 (Bochkovskiy et al., 2020) and, more recently, YOLOv5 (2022).

In this work, the YOLOv5 architecture and its available versions have been used as one of the best-performing methodologies in damage detection and image recognition. The methodology has been applied to a dataset of photos representing the most common bridge defects in the existing infrastructures in Southern Italy.

3. Methodology

In this Section, the application of YOLO has been tested on a specific dataset collected on existing bridges.

3.1. Reference dataset of bridge defects

The reference dataset of bridge defects has been created by collecting figures from the observations of some existing bridges in Italy. The database was composed by 2.685 images of structural elements of bridges (e.g., girders, deck, piles, pile caps), each presenting several damages of different typologies. Each photo was labeled by identifying the structural element typology and observed defects according to the Italian Guidelines (2021). Because of the detailed discretization of the defects proposed, the first response of the labeling was that amount of data was small and not sufficient for running the next steps.

Hence, some defect typologies have been grouped under more generic categories to enlarge the available dataset for the labelling phase. For example, the differentiation of cracks in the reinforced concrete elements among vertical, horizontal, and diagonal was not considered, and all cracks were grouped within a unique category. The final defect categories considered are:

1. Cracks.
2. Corroded steel reinforcement.
3. Deteriorated concrete.
4. Honeycombs.
5. Moisture spots.
6. Pavement degradation.

The six defect typologies are illustrated in Figure 1, and all photos of the dataset have been labeled by domain experts accordingly. Information on the severity and extension of the defects are not specified in this step of the work.

3.2. Object detection

For the object detection task, we identified two possible formulations.

- In the *multiclass* formulation, the object detector is trained over the set of classes previously described. The outcome of this formulation should be a detector able to evaluate whether the defect is located within the picture, and to which class it belongs.
- In the *binary* formulation, the object detector is trained only to evaluate whether a certain patch of the image represents a defect or not. As such, in this formulation the detector is not able to discriminate between different types of defects.

To deal with these problems, we tested the dataset by means of several available versions of the YOLOv5 dataset, ranging from the smallest (i.e., the YOLOv5n) to the medium-sized (i.e., YOLOv5m). Furthermore, tests have been conducted on two revisions of each architecture, i.e., the versions v5 and v6. The main features of all available YOLOv5 architecture are described in Table 1. In this latter, the mAP column reports the mean average precision of each architecture on the COCO 2017 dataset (Lin et al., 2014). It is worth noting that the main difference between versions v5 and v6 architectures lies in the overall number of parameters, which is significantly higher in the v6 revisions. Still, despite a significant increment in terms of the number of parameters, both YOLOv5l and YOLOv5x architectures do not offer a significant improvement in terms of mAP over smaller architectures and then, these architectures have been excluded from the evaluations. In order to compare the reliability of the selected architectures, two different tests have been performed: (a) training from scratch the whole architecture; (b) use transfer learning leveraging predefined weights achieved during training on the COCO 2017 datasets.

Fig. 1. Defect typologies investigated. From the top to down and from left to right, the following defects can be recognized: moisture spot, corroded steel reinforcement, deteriorated concrete; cracks; pavement degradation; honeycombs.

Table 1. Comparison of available YOLOv5 architectures

Model	Model parameters (millions)	mAP [0.95]
YOLOv5n	1.9	28
YOLOv5s	7.2	37
YOLOv5m	21.2	45
YOLOv5l	46.2	49
YOLOv5x	86.7	51
YOLOv5n6	3.2	36
YOLOv5s6	12.6	45
YOLOv5m6	35.7	51
YOLOv5l6	76.8	54
YOLOv5x6	140.7	55

4. Experimental results

When a significant amount of data and computational power are available, training an object detector from scratch is usually the preferred option. Specifically, the direction provided in Lin et al. (2014) states that more than 1.500 images with over 10.000 annotations per class should be used. Still, even if the current dataset does not meet these requirements, we trained the network from scratch to have a baseline performance to which other methods of training (that is, transfer learning) could be compared.

All experiments have been performed using a machine equipped with an Intel Core i9-11900K@3.5 GHz, 32 GBs of RAM and an NVIDIA GeForce RTX 3080 with 10 GBs of onboard RAM. Each network has been trained for 300 epochs, using fixed values for hyperparameters, as described by Lin et al. (2014). Table 2 shows the results in terms of mAP, precision (P) and recall (R), while Figure 2, precision and recall are achieved by training from scratch each architecture over 300 epochs.

From the above results, it is evident that larger models achieve better results. Furthermore, results for the binary formulation are slightly better than results achieved for the multiclass formulation. This suggests that there may be some intra-class visual dependencies which cause an overall instability in the results achieved by the object detector that, as a consequence, achieve suboptimal results. However, there is a general improvement in the mAP when a lower value of IoU is considered. This suggests that lowering this threshold may improve results; however, a lower IoU implies that the network considers in its results also bounding boxes which are only partially overlapped with the labeled patch, which may result in a larger number of false positives.

Table 2. Comparison of performance achieved by different architecture trained from scratch in both formulations

Model	Binary				Multiclass			
	P	R	mAP[0.95]	mAP [0.5]	P	R	mAP[0.95]	mAP [0.5]
YOLOv5n	22.11	18.23	3.14	10.74	30.59	11.39	1.11	3.84
YOLOv5n6	22.52	21.54	3.80	30.59	32.49	12.34	1.79	5.47
YOLOv5s	26.32	21.80	4.70	32.49	20.97	13.42	1.90	5.92
YOLOv5s6	29.15	24.90	5.79	20.97	12.58	18.30	2.83	8.59
YOLOv5m	30.62	26.74	6.56	12.58	16.29	19.25	3.41	9.88
YOLOv5m6	35.53	28.08	7.74	16.29	26.32	18.11	4.84	13.47

As previously stated, the available database is not currently adequate to train a YOLOv5 model from scratch. To overcome this issue, a commonly exploited solution is the use of transfer learning, which allows for a partial retrain of an already trained model. Specifically, in the context of deep neural networks, transfer learning usually involves freezing the weights learned by the initial part of the network, which characterize generic features, whereas latter layers are re-trained on the specific dataset. As for YOLOv5, transfer learning freezes the backbone of the network, which is responsible for extracting the feature maps from the processed image, while the head of the network, which is responsible for object detection, is actually retrained on the specific dataset. Results achieved after 300 training epochs with transfer learning are shown in Table 3.

From Table 3, it is clear how transfer learning outperforms results achieved from the trained-from-scratch architectures. Furthermore, achieved results confirm the evidence that larger models deliver improved values for each metric, and that the multiclass formulation is more challenging with respect to the binary one.

Let us also note that precision is systematically higher than recall, meaning that the networks have a larger number of false negatives with respect to the number of false positives. Visually, this means that the model often completely misses a defect, even if it is more accurate if tasked to correctly find to which class it belongs.

Fig. 2. Comparison of all YOLOv5 architectures trained from scratch in the multi-class scenario.

Table 3. Comparison of performance achieved by different architecture in both formulations using transfer learning

Model	Binary				Multiclass			
	P	R	mAP[0.95]	mAP [0.5]	P	R	mAP[0.95]	mAP [0.5]
YOLOv5n	34.75	28.13	7.27	21.83	27.50	22.20	4.96	14.83
YOLOv5n6	40.10	25.75	7.24	22.63	35.15	21.55	5.80	17.06
YOLOv5s	35.93	27.77	7.81	21.96	37.70	23.11	6.78	18.30
YOLOv5s6	44.96	27.30	8.75	24.12	39.30	22.54	6.82	18.38
YOLOv5m	46.70	27.87	10.01	25.62	35.28	25.41	7.99	18.87
YOLOv5m6	48.81	26.92	10.29	26.18	43.63	24.24	8.34	20.66

5. Conclusions and further works

In this paper, we have presented an initial database of images containing defects in structural elements of existing bridges. The dataset is composed of 2.685 images and is currently under active development. Preliminary analyses performed using the YOLOv5 architecture show some interesting results, which allow to draw some conclusions to be considered in future works. First, the dataset itself appears to be extremely challenging. This is mainly related to two aspects, that is, *inter-class visual similarities*, which may cause a defect of a certain class to be visually misjudged with a defect of a different class (e.g., a corroded steel reinforcement misjudged with deteriorated concrete), and *causal relationships between classes*, where one class of defect causes the occurrence of another and, as a consequence, the two defects are overlapped in the same patch of the image. Second, the dataset is currently highly imbalanced. This causes the model to underperform, achieving suboptimal results.

As a consequence, future works will be focused on four main points. First, the current dataset will be greatly improved by adding several more images taken under varying conditions. Second, data augmentation and balancing techniques will be greatly exploited to deal with the aforementioned issues. Third, larger and more advanced models will be tested, using also improved techniques such as model ensembling, which allows to fuse the results achieved by several models, therefore improving their overall results, and hyperparameters evolution, which allows to select optimal hyperparameters to be used by the model. Fourth, explainable artificial intelligence will be used to analyze the salient parts of the images as provided by the backbone, therefore highlighting the most discriminative traits and providing useful hints on data acquisition and labeling.

Acknowledgements

The first author acknowledges funding by italian ministry of university and research, within the project 'PON - Ricerca e Innovazione - 2014-2020, CODICE CUP (D.M. 10/08/2021, n. 1062): D95F21002140006; (D.M. 25/06/2021, n. 737): D95F21002160001'. Moreover, authors acknowledge the FABRE Consortium for the financial support, within the agreement named "*Supporto tecnico-scientifico per lo sviluppo della metodologia per il censimento, ispezioni iniziali e individuazione delle Classi di Attenzione di un campione di ponti e viadotti gestiti da Anas Spa; prioritarizzazione delle operazioni di valutazione di livello 4; verifica della qualità e di omogeneità dei risultati*".

References

Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M., 2020. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934

Cardellicchio, A., Ruggieri, S., Leggieri, V., Uva, G., 2022. View VULMA: Data Set for Training a Machine-Learning Tool for a Fast Vulnerability Analysis of Existing Buildings. Data. 7(1):4. https://doi.org/10.3390/data7010004

Cha, Y.J., Choi, W., Büyüköztürk, O., 2017. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. Computer-Aided Civil and Infrastructure Engineering, 32:361–78. https://doi.org/10.1111/mice.12263.

Cha, Y.J., Choi, W., Suh, G., Mahmoudkhani, S., Büyüköztürk, O., 2018. Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types. Computer-Aided Civil and Infrastructure Engineering, 33:731–47. https://doi.org/10.1111/mice.12334.

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587.

https://github.com/ultralytics/yolov5, last fetched on 8th March 2022.

Li, R., Yuan, Y., Zhang, W., Yuan, Y., 2018. Unified vision-based methodology for simultaneous concrete defect detection and geolocalization. Computer-Aided Civil and Infrastructure Engineering, 33(7), 527-544

Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C.L., 2014. Microsoft coco: Common objects in context. In European conference on computer vision, pp. 740-755, Springer, Cham.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C., 2016. Ssd: Single shot multibox detector. In European conference on computer vision, pp. 21-37, Springer, Cham.

Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., Omata, H., 2018. Road damage detection and classification using deep neural networks with smartphone images. Computer-Aided Civil and Infrastructure Engineering, 33(12), 1127-1141.

Ministero delle Infrastrutture e della Mobilità sostenibili. Decreto ministeriale numero 493 del 03/12/2021: Linee guida per la gestione del rischio dei ponti esistenti e delle Istruzioni operative per l'applicazione delle Linee Guida stesse. 2021. (in Italian)

Potenza, F., Rinaldi, C., Ottaviano, E., Gattulli, V., 2020. A robotics and computer-aided procedure for defect evaluation in bridge inspection. Journal of Civil Structural Health Monitoring, 10(3), 471-484.

Prasanna, P., Dana, K.J., Gucunski, N., Basily, B.B., La, H.M., Lim, R.S., 2016. Automated Crack Detection on Concrete Bridges. IEEE IEEE transactions on automation science and engineering. https://doi.org/10.1109/TASE.2014.2354314.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788.

Redmon, J., Farhadi, A., 2017. YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7263-7271.

Redmon, J., Farhadi, A., 2018. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28, 91-99.

Ruggieri, S., Cardellicchio, A., Leggieri, V., Uva, G., 2021a. Machine-learning based vulnerability analysis of existing buildings. Automation in Construction, Volume 132, 103936. https://doi.org/10.1016/j.autcon.2021.103936

Sun, H., Burton, H.V., Huang, H., 2020. Machine learning applications for building structural design and performance assessment: state-of-the-art review. Journal of Building Engineering, 101816.https://doi.org/10.1016/j.jobe.2020.101816.

Xie, Y., Ebad Sichani, M., Padgett, J.E., DesRoches, R., 2020. The promise of implementing machine learning in earthquake engineering: a state-of-the-art review, Earthquake Spectra 36 (4) 1769–1801. https://doi.org/10.1177/8755293020919419.

Xu, Y., Bao, Y., Chen, J., Zuo, W., Li, H., 2019. Surface fatigue crack identification in steel box girder of bridges by a deep fusion convolutional neural network based on consumer-grade camera images. Structural Health Monitoring. https://doi.org/10.1177/1475921718764873.

Yang, X., Li, H., Yu, Y., Luo, X., Huang, T., Yang, X., 2018. Automatic pixel-level crack detection and measurement using fully convolutional network. Computer-Aided Civil and Infrastructure Engineering, 33(12), 1090-1109.

Zhang, A., Wang, K.C., Li, B., Yang, E., Dai, X., Peng, Y., Chen, C., 2017. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network. Computer-Aided Civil and Infrastructure Engineering, 32(10), 805-819.

Zhu, J., Zhang, C., Qi, H., Lu, Z., 2020. Vision-based defects detection for bridges using transfer learning and convolutional neural networks. Structure and Infrastructure Engineering, 16(7), 1037-1049.