ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 12 (2022) No. 6, pp. 163-191
DOI:10.33168/JSMS.2022.0611

Identify Faults in Road Structure Zones with Deep Learning

Yohan Roy Alarcón Cajas, Yenmy Zúñiga Guisado, Afredo Daza Vergaray

Universidad César Vallejo, Carretera Panamericana Norte Km 695, Los Olivos, Lima, Perú

yalarconca@ucvvirtual.edu.pe, yzunigag@ucvvirtual.edu.pe, adazave@ucvvirtual.edu.pe

Abstract. Traffic accidents occur due to the deterioration of road structures, this poor condition can lie in the low quality of its components, climatic changes, seismic zones, heavy cargo transport traffic, among others; the purpose of this research is the detection of pavement faults. The research was carried out taking into account the following stages: collection and preparation of data that consisted of the capture of images with a smartphone in the city of Lima-Peru, resizing and labeling, architecture network where the training process was captured, processing that consists of the entry of the algorithm and the deployment that is the implementation of the same in a system to detect cracks and gaps; for which the YOLOv5 algorithm was used, where for the training of the algorithm 420 images with 100 epochs were used, having an accuracy of 58%, and then validated with 30 images, having the averages of 85% accuracy, 91.5 sensitivity and 81.5% F1-Score. It is concluded that the algorithm with convolutional neural networks helps to properly identify pavement failures so that the authorities can do the proper maintenance.

Keywords: deep learning, pavement, cracks and voids

1. Introduction

The road structures must be in good condition, because through its main arteries transit commerce, the intercommunication of its inhabitants and exchange of different cultures; it should be noted that the construction of road structures aims at the development and economic growth of a country (Shatnawi, 2018). It is necessary to mitigate the defects that may be generated over time as a result of the weather or transport of heavy load; therefore, the importance of carrying out constant monitoring that keeps the useful life cycle of road structures in good condition, using concrete pavements continuously and being able to mitigate the risks of accidents that they may cause (Shim et al., 2020).

Defects in the roads are caused by different reasons, where car accidents stand out, which weakens their components and causes dangerous areas for road traffic (Chun & Ryu, 2019), also known as the disease of road structures (Zhang et al., 2020); aging is another factor that weakens its structural components, which are weakened and whose state must be evaluated periodically to identify these failures (cracks or gaps) and thus implement an adequate rehabilitation with quality components (Slami & Yun, 2021), it is important that the State or competent authorities intervene in the analysis and corresponding evaluations for the adequate transit of society.

In many underdeveloped countries, road infrastructure maintenance activities have declined due to the complexity of access to road infrastructure, bureaucracy or other factors (Silva et al., 2020).

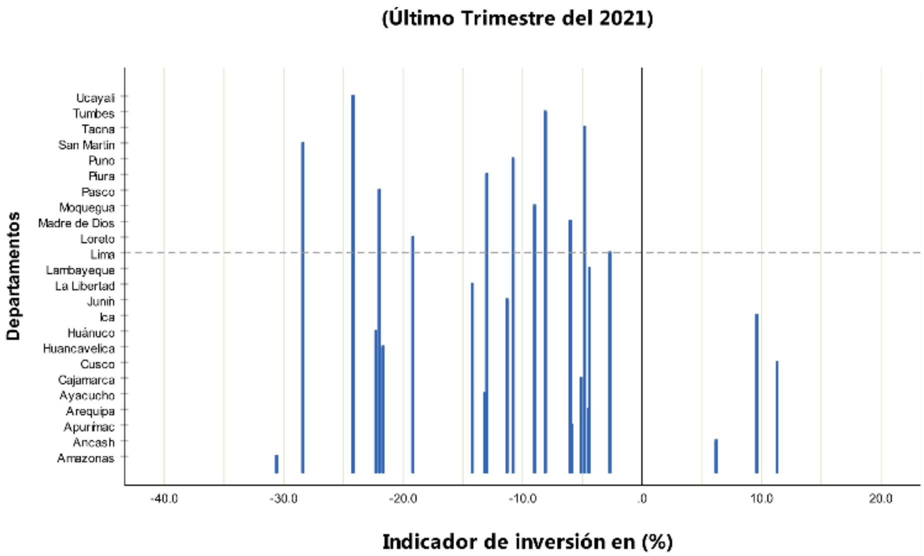


Fig. 1: Registration of the investment index that is applied in the construction sector according to its geographical area of Peru (INEI, 2022)

Figure 1 shows investment in the construction sector according to its geographical area. So much so that, in the last quarter of 2021, investment in structures in Peru was -5.8%, while in the departments of Amazonas highlights indicators of -30.6%, unlike Ancash where it registers positive percentages of 6.2%, Apurímac is below -5.9%, which is at -4.5% in Arequipa, while in Ayacucho it shows indicators of -13.2% of investment, also Cajamarca shows -5.1%, and Cusco acquires high percentages such as 11.3% in execution of road investments, on the other hand the department of Huancavelica registers -21.7% unlike Huánuco that registered -22.3%, compared to Ica where only 9.6% appears, Junín with 11.3%, La Libertad with -14.2%, Lambayeque indicators of -4.4%, Lima only executed the investment in construction of -2.7% which is a very low indicator in this sector, while Loreto shows an execution of -19.2%, Madre de Dios with -6.0%, Moquegua shows -9.0%, Pasco with -22%, Piura with a record of -13%, Puno -10.8%, San Martin with -28.4%, Tacna -4.8%, Tumbes with -8.1% and Ucayali -24.2% (National Institute of Statistics and Informatics [INEI], 2022).

The budget in the construction sector is composed of investments in road structures and their rehabilitation; in addition, it is important to intervene in technological tools that allow standardizing the identification and analysis of the failures generated in these structures. Therefore, it is important to mention the benefits of technology in this area, specifically artificial intelligence has been involved in the personal, work and economic activities of man (Backman & Chung, 2021; Opara et al., 2021); points out that image recognition through deep learning is achieving important developments for society, such as the detection of objects with a high degree of accuracy (Qiao et al., 2021); where neural networks, algorithm architectures, techniques and effective tools are used, which have the ability to learn and recognize patterns to make decisions (Sarmiento, 2020; Coluccia, 2021; Morales & Kayacan, 2020).

This article is aimed at detecting damage to road structures such as gaps and cracks, caused by various reasons, such as the intervention of the human factor, climatic changes, seismic zones, heavy cargo transport transit, which leads to an imminent danger for the population in transport. Likewise, to identify and analyze these failures, it was necessary to carry out a series of activities such as capturing images through a smartphone that allows us autonomy in data collection, choosing algorithms, training the intelligent model and measuring the results (Morales & Kayacan, 2020; Janani et al., 2022). In this context, the research work is related to identifying faults in the road structures in which image processing techniques are applied through convolutional neural networks (CNN) (Kumar et al., 2021; Zhang, 2021). To do this, it is required to perform a series of structured steps where algorithms are incorporated that provide data for classification and take this information from the fault or defect. It is worth mentioning that tests have been carried out with a group of images taken from the population group being the positive results in the determination of the accuracy for the identification of faults in the road structures (He & Liu, 2020; Enzo, 2021).

The rest of this document is planned as follows. Section 2 summarizes some related work. Section 3 explains the proposed methodology. Section 4 illustrates the results obtained. The above results are discussed in section 5. Finally, section 6 concludes the document.

2. Literature Review

The works developed to identify cracks, as well as gaps through Deep Learning consist of authors who obtained highly qualified results with CNN, applied in different studies in order to automatically identify faults of road pavements based on intelligent architectures through the segmentation of images; this must include a series of algorithms that served as gears to obtain good accuracy (Fan et al. 2022).

However, it should be deepened that, in the real field, the inspection of the roads is carried out with topographic equipment, which must be clear for its adequate treatment by specialized agencies, but for image recognition investigations there is no standardization of data that are available, some studies took the method of image collection through smartphones as mentioned (Maeda et al. 2018), while in other researches the collection of information was supported by a drone (Kumar et al., 2021).

The various ways of grouping information for image segmentation are essential in the application of CNNs through architectures such as YOLO-v3, among the most used. This algorithm was used for the detection of cracks in pavements in three phases that included image processing, machine learning and deep learning. More detail on the algorithms and methods used is provided below.

2.1.1. Traditional Methods

Traditional methods work with the finding of pixels in images, containing shapes of cracks or gaps, assigning grayer pixels; There are also other methods of detecting cracks or gaps that use global or local threshold techniques to segment the anomalies of these pathways through strokes (Zhang, 2021). However, it is worth mentioning that these methods are very weak to noise, since the conditions of an image vary by the quality or optical position; that is why there is the complexity of selecting the appropriate threshold, under these concepts there are other techniques that allow the traceability of the edges of the cracks or holes; but when the quality of the contrast has low indicators, the sharpness of the target will be weakened, complicating the detection of road diseases. Therefore, combining the different pixel segmentation techniques, making use of deep coding, will allow optimal indicators to be obtained (Augustauskas & Lipnickas, 2020).

2.1.2. Machine Learning

This method became an important form within artificial intelligence, especially to detect cracks in pavements more accurately; so much so that in recent years they have been used on a large scale for image detection processing, with which they obtained highly reliable results (Augustauskas & Lipnickas, 2020). Likewise, these forms of work run under a segmentation approach of classified images (Mukhopadhyay et al., 2021), as well as those supervised with multiple layers or processes, allowing greater robustness (Zhang, 2021; Xiao et al., 2021).

2.1.3. Deep Learning

The methods applied in deep learning were born in the 1986s because it integrates CONVOLUTIONAL ARCHITECTURES CNN (Convolutional Neural Network), thanks to the solid increase of information, which allowed to design and program new sophisticated image processing units with complex mathematical formulas, such as vectors or matrices found in the new models of Artificial Intelligence that increases over time.

In that sense, Deep Learning facilitated machine automation tasks, this is a new form of learning, which is based on continuous or multilayer layers that are presented in neural networks. However, these types of architecture can reach indicators higher than 0.5 and rise to their greatest study in the use of these applications developed through the technology industry.

In the 2015s, free writing libraries such as TensorFlow, PyTorch for the development of machine learning began to be applied in this industry (Abeliuk & Gutierrez, 2021). According to Mukhopadhyay et al. (2021) it guarantees the quality in deep learning applied in the identification of cracks and gaps on road structures, which makes this structure reliable.

In addition, to work with deep learning technology there are different methodologies such as the general adversarial cycle network (GAN) model (Kim et al., 2022), convolutional layers merged with LSTM, convolutional layers merged with GRU and CNN (Ghaleb et al., 2022).

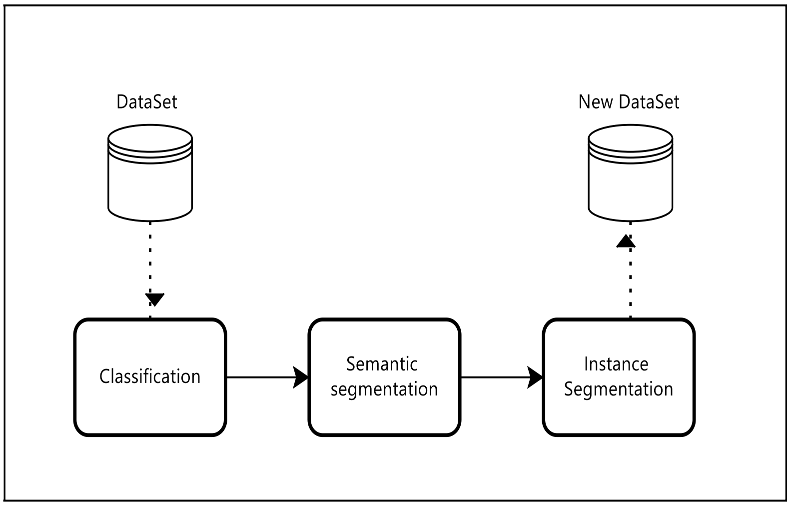


Fig. 2: Transition of THE CNN in its classification and segmentation of images

Figure 2 shows how CNNs work to classify, as well as segment the images semantically and by cases, and then transform them into a new set of data on which to work. Likewise, in the segmentation, the performance behavior of the applied method was evaluated, in order to compare the operation of the segmented neural networks, as well as the neural multiscale type to homologate a standard, for this only the labeling technique is used and precision data is obtained (Shim et al., 2020).

3. Methodology

It is important to detail the process of data collection and preparation that is required in the processing of images through Deep Learning techniques, which will be implemented with sophisticated algorithms such as YOLOv5, as well as the construction of a system that has the ability to recognize images in the execution of the CNN convolutional neural network model, which will allow to detect cracks and holes on asphalt pavement or concrete structures. Likewise, the image recognition system was built through image capture, image classification and resizing, creation of the dataset_clasificadorGH, training, acquisition of the new model, development of the system and implementation of the new dataset model.

3.1. Data collection and preparation

3.1.1. Data collection

For the development of the present work, the process of capturing 420 images in RGB format was carried out through a smartphone, for this the position of the camera was taken at an angle between 15° and 45° due to the variety of positioning of captures of the transit routes that contains forms of cracks and holes, which are stored in the memory of 32Gb images with an original size of 2160 x 3120 pixels.

Table 1: Collection and recording of image capture features

Class	Defect	Type of pavement	Number of images	Total by class
Crack	Transverse crack	Cement	121	292
		Asphalting	12	
	Longitudinal crack	Cement	84	
		Asphalting	8	
	Crocodile crack	Cement	33	
		Asphalting	34	
Hollow	Well cover	Cement	50	128
		Asphalting	78	

Table 1 shows the collection and recording of the sample classified by the type of crack and hole class, according to its type of defect, type of pavement (cement or assault) and number of images, to determine the performance of the result in detection; it is important to collect several captures of images containing cracks and gaps, to use in training and to be able to detect the type of class in the result. The images may vary, depending on the area of study that the information collection is carried out, it will become part of the dataset.

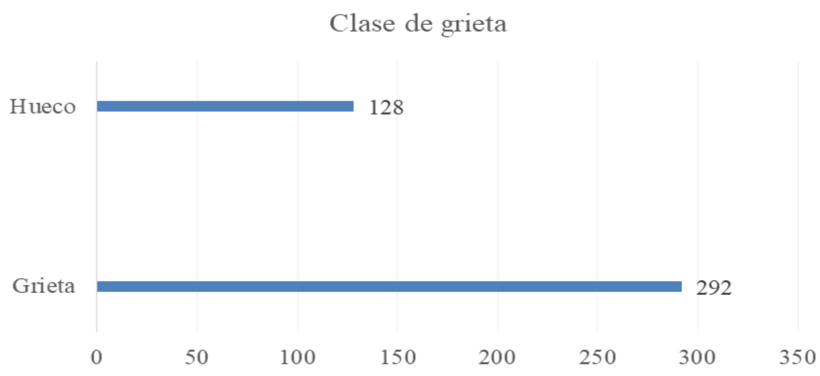


Fig. 3: Registration of images classified by class

In graph 3, the total of the images collected by class is shown, appreciating an indication of 292 in cracks and 128 in holes, a fact that allows a model to detect cracks, without neglecting the detection of holes in asphalt or concrete structures.

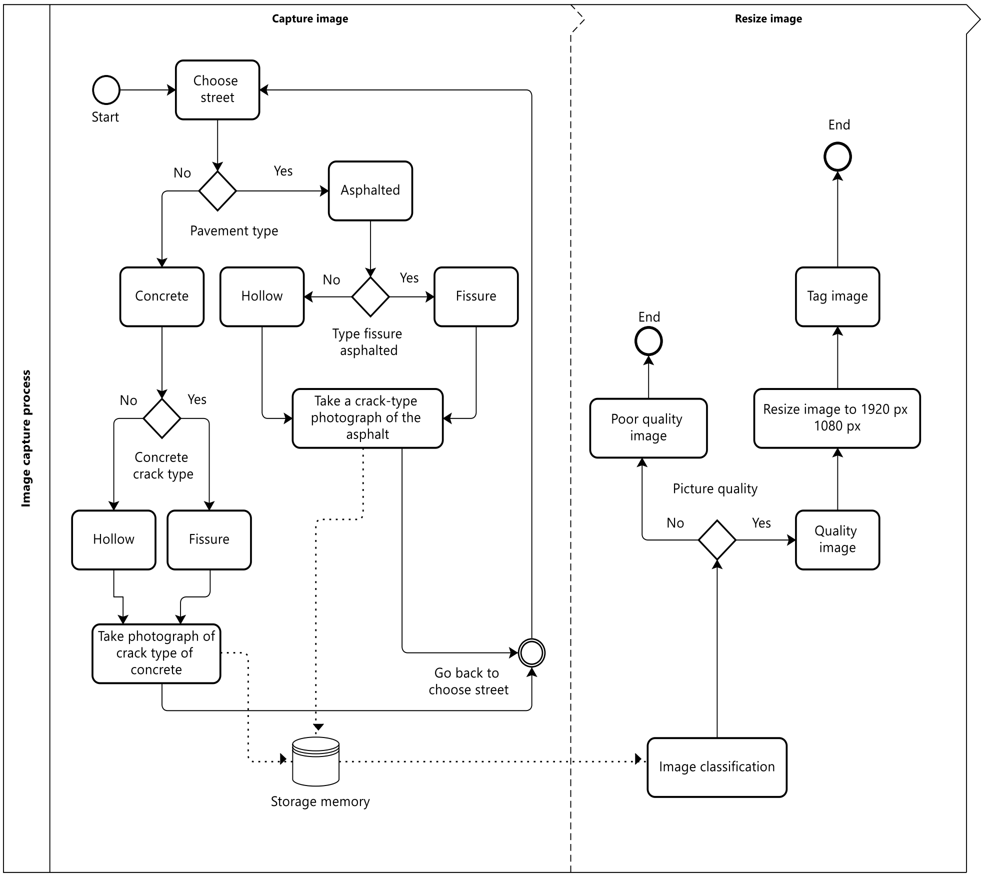


Fig. 4: Process diagram applied in image capture

Figure 4 shows the flow that was developed for the collection of images, at this stage the details that were made are mentioned, in principle where an analysis of the geographical area was made for the acquisition of the images by means of a smartphone taking those avenues or streets that contain characteristics of cracks and gaps on asphalt or concrete pavement structures, this process consists of two phases:

The first phase, called image capture, the study area or street is chosen, then the type of pavement is verified, according to the first two characteristics we proceed with the image capture, where the type of class (crack or hole) culminated with the capture is visualized, the information is stored in the memory of the smartphone and the process is restarted in order to capture another image in a way successive in the process.

The second phase, called image resizing, consists of the classification of the captured images that contain good resolution of pixels and to comply with this condition proceeds to the resizing of the images to 1920 pixels by 1080 pixels, through a technological design tool, this process is essential to standardize the dimensions of height and width captured with the smartphone and reduce the size of each image, these features will optimize the calculation that is made during the training of the images through deep learning techniques.

3.1.2. Data preparation.

Likewise, the classification in sectors is carried out to improve the performance of the convolutional architectures during their training, that is why in the first instance the images are grouped according to the dimension, this technique allows to improve the mathematical calculations of the receptive capacity of the process, through which it emits the results of sample segmented into pixels, the image base was set to a size of 1920 x 1080 pixels to maintain sharpness quality, to contribute to the process normalization functions and the impulses of the Adam algorithm optimizer, as well as the intervention of max pooling, which allows to group the parameters of height, width, type and length of the RGB format of the image, with the first execution that occurs in a certain layer, this event is important for the performance during the execution of the epochs, important data for the training of the Convolutional Neural Networks CNN.

All the information collected on the pavement structure was acquired in a region of Lima, Peru, as well as the development of the process mentioned in Figure 4, in which seven streets were considered to analyze cracks and voids, including shaded tracks, wet, small particles of earth or dust on structures with cracks, paint stains, different schedules, varied climate between 17°C and 25°C, characteristics that are important for training since these will strengthen the recognition of the patterns that will be responsible for detecting different shapes or moments.

Table 2: Collection of images in different states of cracks and holes.

	Original (4160x3120 px)	Resized image (1920x1080 px)
Crack, on concrete with paint		
Crack on concrete		
Crack on asphalt with shade		

Crack on asphalt with moisture inclusion		
Hollow on asphalt, including moisture		
Hollow on asphalt		

Table 2 shows some characteristics of the cracks and gaps at the time of data collection; among the most outstanding characteristics are the type of material (concrete or asphalt) and conditions (paint, shade or humidity). Likewise, the dimension of the images was 4160x3120 pixels, which responded to the characteristics of the camera of the mobile equipment used; however, in order to improve the quality of the analysis, they were resized to a size of 1920x1080 pixels.

At this level, the geographical geolocation of the study area was used, with characteristics of cracks and gaps, in order to apply techniques on deep learning, developed in the field of artificial intelligence, as well as sophisticated algorithms that are able to identify objects in very short times and accurately (Opara et al., 2021),, YOLO was implemented, which allows to extract characteristics of the images using internal Darknet algorithm, for the segmentation of data and these are used by convolutional neural networks (DCNN) (Madasamy et al., 2021).

3.2. Architecture Network

On the development of the semantic segmentation architecture and the modules of the CNNs that have been developed during the training, masks of recurrent convolutional neural networks are added to optimize the process, that is, the obtaining of the characteristics of the lower layer are acquired through semantic segmentation that adopts the characteristics of cracks or gaps, in this sense, the phase of the training process of the YOLO algorithmic architecture is applied to the images acquired during the data collection process (Opara et al., 2021; He & Liu, 2020), in addition, Srivastava et al., (2021) makes a comparison on the versions of this architecture and argues that the YOLOv3 version segments through multiple custom layers, thanks to its soft-max binary function, this allows a reduction in the process to mitigate complexity, however the latest model of YOLOv5 is the most sophisticated in the field of performance and execution time, under these characteristics, in the following graphs the training process is shown (Fig. 6) and the difference of YOLO in its different models (Table 3).

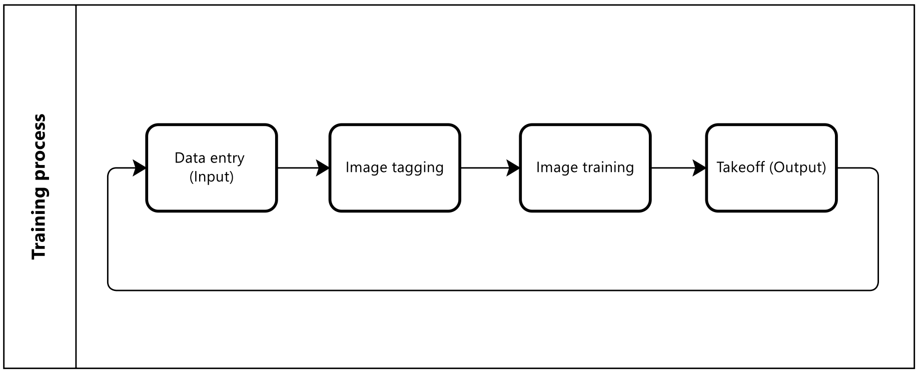


Fig. 6: Training process in YOLOv5 architecture, to implement CNN models

Figure 6 shows the training process that consists of data collection, labeling, training and takeoff of the model. For this it is important to analyze the different models that exist in YOLO, because it varies the performance between its versions, as in the speed of processing layers, as shown in the following table.

Table 3: YOLO algorithms and their performance in different models in COCO dataset.

Nano YOLOv5n	Small YOLOv5s	Medium YOLOv5m	Large YOLOv5I	XLarge YOLOv5x
4 MB	14 MB	41 MB	89 MB	166 MB
6.3 ms	6.4 ms	8.2 ms	10.1 ms	12.1 ms
28.4 mAP	37.2 mAP	45.2 mAP	48.8 mAP	50.7 mAP

Table 3 shows that the sizes vary between 4MB to 166MB, as well as the values in the mean of the precision that is composed between 28.4 mAP to 50.7 mAP, and the performance in milliseconds also varies within the range between 6.3ms and 12.1ms, as well as, the values applied in the present study were 166MB, 50.7mAP, with a process time of 12.1 milliseconds.

Likewise, there are different versions of Yolo among which are v5n, v5s, v5m, v5I, v5x; in addition, each one has its degree of indicator that differentiates the complexity of the neural networks constituted among themselves; under these terms of version improvement, in the present research work the 5x version was applied, due to the high degree of accuracy in the average, which exceeds 50%, being the most appropriate for processing large batches and complex matrix calculations.

Under this concept, YOLO in its latest model allows to work with complex custom datasets, however, developed works were found that used a dataset such as COCO, DarkNet-53, that is, with 53 layers of convolutional neural networks, improving accuracy and speed during the process of object detection, also these architectures are implemented in many studies that contribute to solving social problems (Mukhopadhyay et al., 2021; Opara et al., 2021; Kumar et al., 2021; He & Liu, 2020).

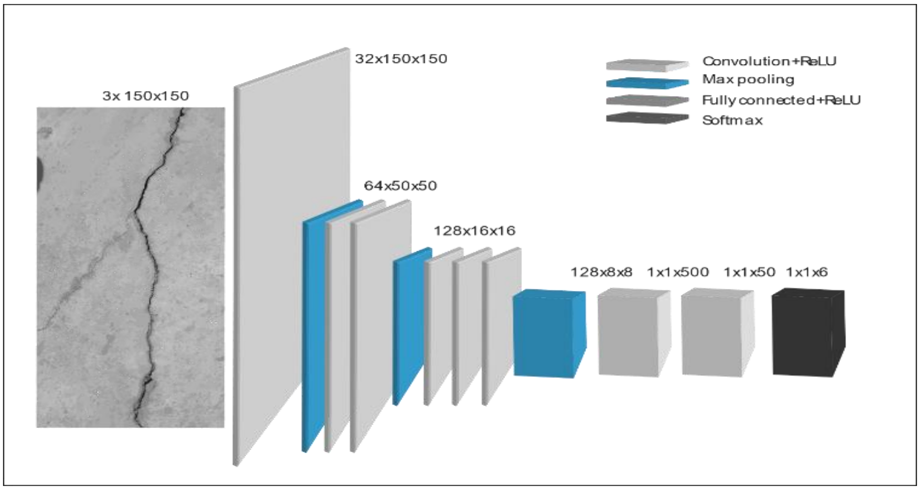


Fig. 7: Adaptation of the structured multilayer convolutional neural network model (Ramalingam et al. 2021; Sun et al. 2021)

Figure 7 presents an adaptation of the YOLO model used for this article, in the first stage the image of the crack or hole is entered, and then it is processed with the convolutional neural networks, applying max pooling until an optimal analysis is achieved.

Therefore, the performance of YOLO is highly reliable thanks to the architecture of its convolutional neural networks that are linked together, this information process is important to obtain the degree of confidence, the ROC curve and the predictive values.

3.3. Processing

The neural network from its structure, has the ability to process information in large batches, thanks to the use of different algorithms that allow to work optimally such as U-Net models, ResUNet (Kittipongdaja & Siriborvornratanakul, 2022), which are used in different studies applied with YOLO algorithms, in this work we will implement a series of activities that are described in the following lines.

3.3.1. Tools for modeling

For the classification, the computational architecture detection model was applied, which is developed written in the Python programming language, free container libraries such as OpenCV, PyTorch were also used. In addition, the labeling was carried out the online service of the website makesense.ai (Skalski, 2022), it is important to know about these technological tools, as well as the libraries that will be incorporated into the intelligent system that is able to recognize cracks and gaps that are in asphalt pavement structures or in structures composed of concrete.

Python is widely used in the technology industry to develop versatile software, due to its robustness and free license. that offers more than a set of useful libraries in the development of scientific research works and other relevant fields (Van Rossum, 2022), under that premise this language was made by Guido Van Rossum in 1989, being the first programming language most used in the software technology market, according to the report of surveys published by the prestigious company Tiobe Software (Jansen, 2022), in which I collect information on the use of the most active programming languages until June 2022 (Cass, 2021), which guarantees an intelligent system that will be developed with Python in its version 3.9.

PyTorch, is one of the fundamental components for the design of neural networks, since it offers a range of free libraries and composed of a set of tensors that are structured in a multidimensional way, which are specially designed to optimize and implement in the present research work, thanks to the deep learning techniques offered by artificial intelligence. Likewise, this mechanism of architectures is widely implemented in projects of the recognition and detection of classes, on pavement structures such as cracks and voids, making use of convolutional neural network algorithms (Arbaoui et al., 2021; Shim et al., 2020), very sophisticated in the precision and speed of the process (Qiao et al., 2021).

Numpy, is a fundamental library that is incorporated in Python (Oliphant, 2022) the same that provides information in multifunctional matrix, as well as masked matrices, which are essential for studies related to artificial intelligence, under that concept this library is very useful when working with image labeling, which must be transformed into matrices so that the algorithms of YOLOv5 and others can read the content, to predict or identify objects, all these functions are written under mathematical, logical and algebraic functions, as well as operations related to statistics.

Colab, is a free service offered by the company Google, as well as geolocation APIs, for which as a first step we will use Google Colaboratory-Colab, where it provides virtual machines with graphic processing units including RAM, which streamlines the processes developed under deep learning, optimizing the training of neural networks, also does not recharge the processor.

Qt, is formed by a set of libraries written in C ++, compatible with multiple operating systems, which allows to be a high-level library, accessible in many desktop, mobile, web and geolocation devices (Qt Group, 2022), also, as a development of custom user interface environments, allowing the development of quality products, then Qt has used the PyQt5 library, to be able to link Python and Qt version 5, being the complete framework for the development of system interface.

3.3.2. Model training

To start with the training process, CNN models are applied for regression and classification, the areas of cracks or gaps that are required to be obtained as a result in pixels are recorded and framed more accurately, this process is also called multiclass because in an image we can find a variety of classes interposed one on top of another, which makes the matrix process more complex.

Once the image labeling process is finished, the generated dataset is incorporated into The Colab services, for the training of the YOLO algorithm in its v5x version, in order to detect cracks and gaps in the structural surfaces of asphalted or concrete pavements, through the use of methods applied in CNN convolutional neural networks.

3.4. Deployment

3.4.1. Tools for the system

To develop the interface of the system, technological tools were used such as PyCharm Community Edition in its version 2022.1.2, Spyder in its version 4, Anaconda, a webcam device and Adobe Inc., which was applied for vectorization, resizing, creation of geometric figures in the application and having a quality product, in the visual aesthetic field for the user; Also, it is complemented by the use of the Visual Paradigm software in its version 16.3, where the different processes for the development of the application were structured, as well as Bizagui Modeler, Archi – archimate modelling in its version 4.8.1, which complements with the design and workflow of the systemic work.

Classifying the technological tools, we proceed with the writing of the code under the Python programming language, and the incorporation of the dataset model trained under the YOLOv5 algorithm.

3.4.2. Classification of the trained model

This stage is executed after having carried out the training of the model, for the identification of cracks and gaps of road structures; in the objective of the present study, it should not only be to accurately frame the shapes and sizes, but also the false positives, in order to differentiate the framed data.

4. Results

For the techniques applied in the model, the crack and hole data was analyzed, which contains the custom dataset called dataset_clasificadorGH, which has a pre-trained dataset and for validation. In addition, to verify the performance of the model, a test test was carried out using a camera, which included free libraries such as PyTorch, Python, Cv2, pandas, yaml, which were executed in Google Colab, and for tests on a CPU (Central Processing Unit) Intel(R) Core (TM) i7-10700 (2.90GHz) RAM 16GB with Windows 10 operating system based on (64 bits).

As a result of the training, the precision, recall, accuracy and F1 were obtained, calculating the function of the true positives called in its abbreviation (TP), true negatives (TN), false positives (FP) and false negatives (FN), in the following structural way.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

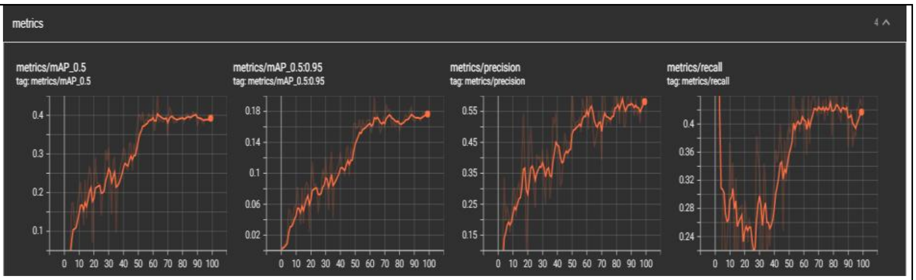
$$F1 = \frac{2xPrecision\ x\ Recall}{Precision + Recall} \quad (3)$$

The precision formulas (1), recall (2) and F1-Score (3) are presented (Opara et al., 2021; Kumar et al., 2021; Mukhopadhyay et al., 2021; He & Liu, 2020) that will be used to measure the functionality of the model.

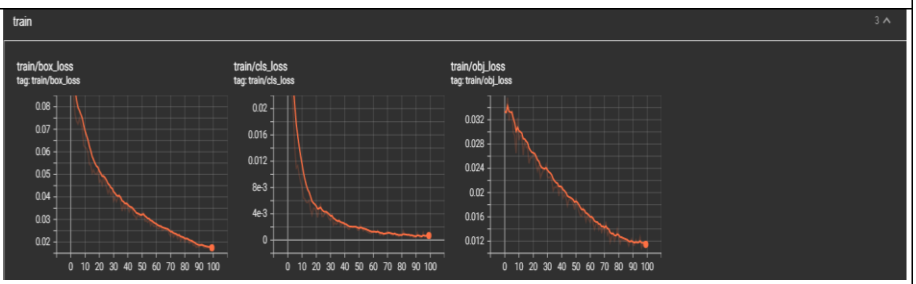
4.1. Training Results

The results of the YOLOv5 training, using Colab and libraries such as Pytorch, Numpy and CV2, allowed to have the following metrics for the identification of screams and gaps on asphalt and concrete structures.

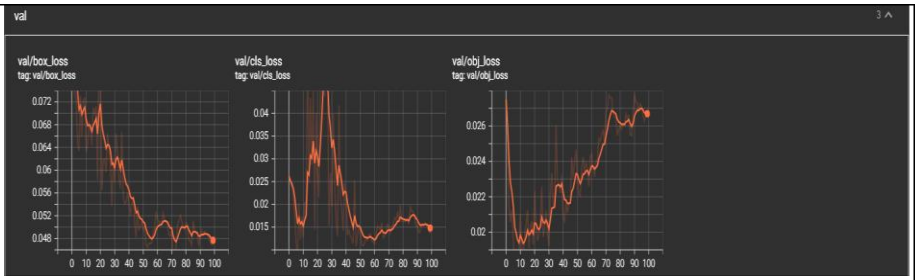
a)



b)



c)



d)

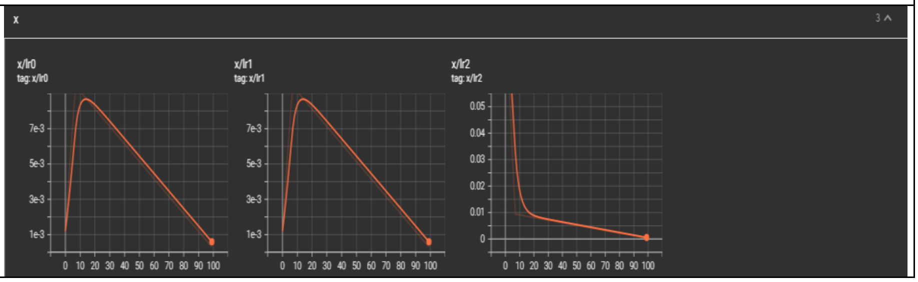


Fig. 8: Dataset image training metrics

Figure 8 shows 4 sections; section a) details the metrics obtained from precision training through semantic segmentation, as well as the mean accuracy of the mean (mAP) of the model that allowed to evaluate the performance of deep learning that is related to accuracy and recovery rate in order to recover the values and plot the precision curve on the x axes, y; this mechanism allows to calculate the average value of all classes, that is, the higher the value in mAP, the better the performance of the model that was applied in 100 epochs, it is important to mention that this value can be improved if more times and an adequate batch or weight of information is assigned. It is also important to mention the margin of error of 0.5 which complements with the margin of acceptability 0.95. While in section b) the training record is appreciated by graphing the progress of the behavior in 100 epochs, so much so that the more times the response time is optimized for the model. Section (c) shows the validation record of the images obtained during training. Finally, section d) represents the bias and weight of the neural network that occurred in 100 assigned epochs.

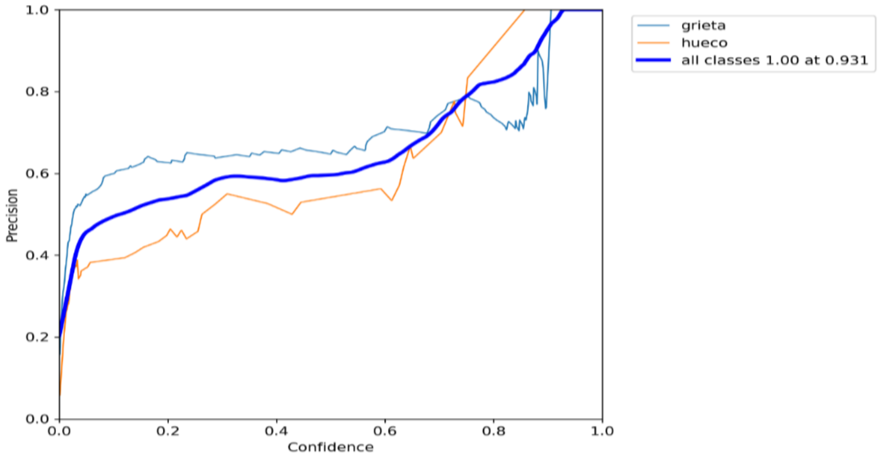


Fig. 9: Accuracy and confidence

As seen in Figure 9, the mean of the precision of the classes (crack - hole) reach values that exceed 0.75 in their reliability being 0.931, which indicates an optimal behavior in the results obtained if there is a crack or hole.

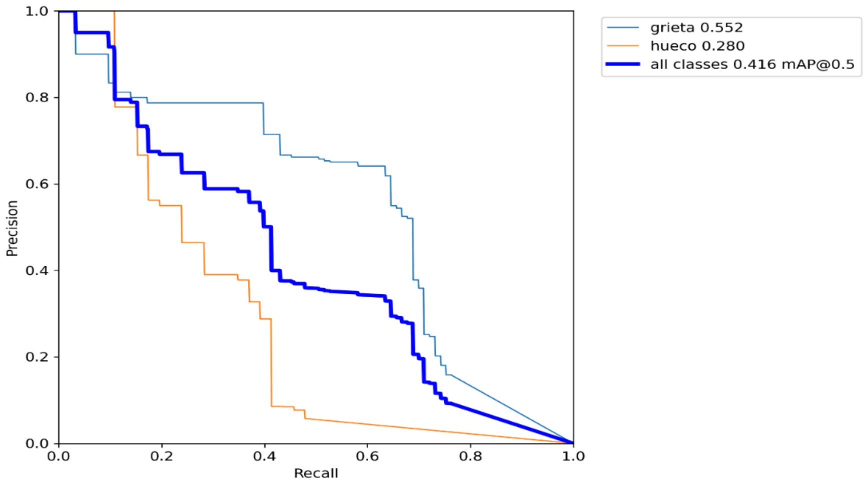


Fig. 10: Accuracy during information retrieval on the ROC curve

Figure 10 shows the average average accuracy (mAP) that begins in the values of the recall where the performance of the model that performs in recovering the information and detection of the classes called cracks and gaps is shown, where values of 0.552 were obtained for the cracks and 0.280 for the hollow class this happens during training with the increasing times, therefore the high average accuracy rate happens more in the detection of crack values that can change by the number of images containing cracks and gaps. While mAP@0.5 represents the union of values in 0.5 of the recall and accuracy.

Likewise, this mechanism represents the behavior of a neural network that retrieves or captures the information of a process as a final result by a (TP) which conforms to the softmax threshold that has the function of emitting the final results of a process.

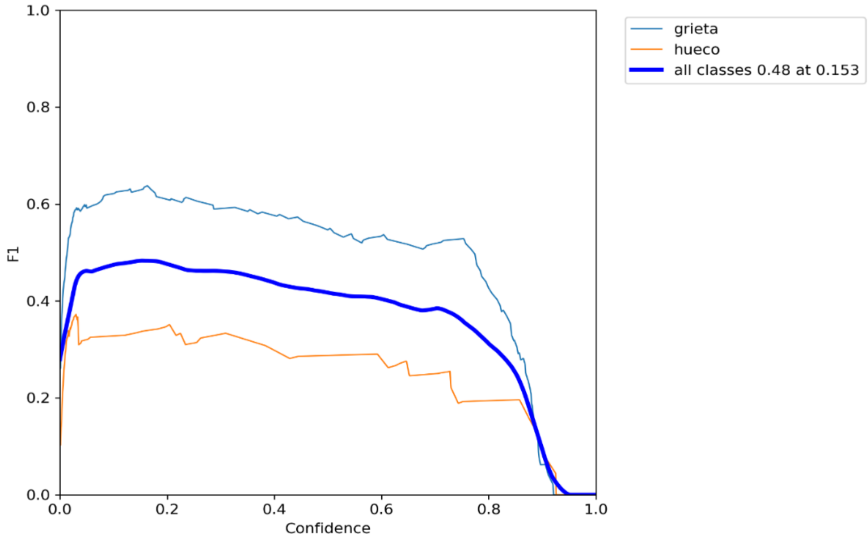


Fig. 11: Curve (F1) during training confidence from 100 eras

Figure 11 shows the score of the class called crack (celestial line) and the hollow class (orange line) as well as the mean between both classes (blue line), since this value is obtained from the results of the precision and the recall averaging the precision in decimal values of the score that will allow to draw the behavior of these classes that occurs on the x axis represented by the confidence and the y axis represented by the F1 creating shapes during the transition occurred in the epochs, this mechanism allows to obtain decimal values that are generated between 0 and 1 in order to validate the reliability of the model applied in the detection of cracks and gaps.

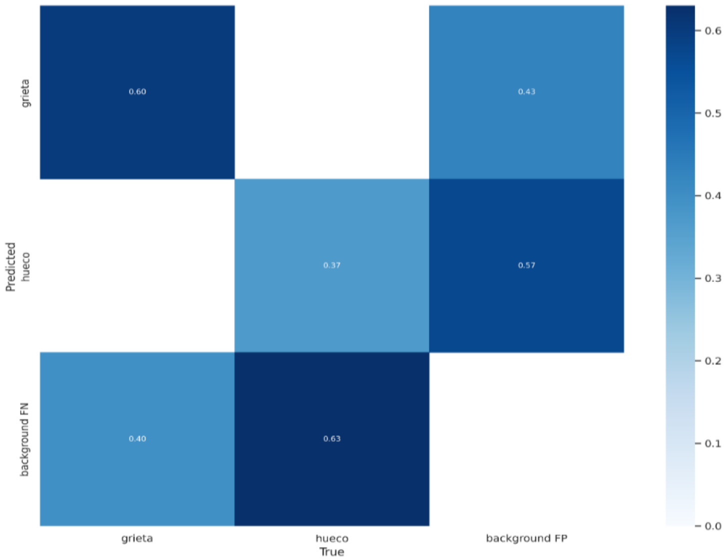


Fig. 12: Confusion matrix

Figure 12 shows the confusion matrix, the aforementioned matrix is composed of a set of information in percentages that represents the successes of the characteristics and criterion of the qualification of the classes that are divided by a cut-off point being fundamental for the classification of the arguments called negative (False) and positive (True) that lead to the union with logical operators and parameterized in values from 0 to 1 which will be used by the Log-loss or called logarithmic loss that occurs during classification.

Also, this indicator indicates that if its value on the label is 0 represents a type of crack class while the value is 1 it will be representing a class of hollow type, these indicators go with approximations of a probability of 0.51 that corresponds to a classification however the margin of error that is in the cut-off point less than 0.5 will indicate a margin of error.

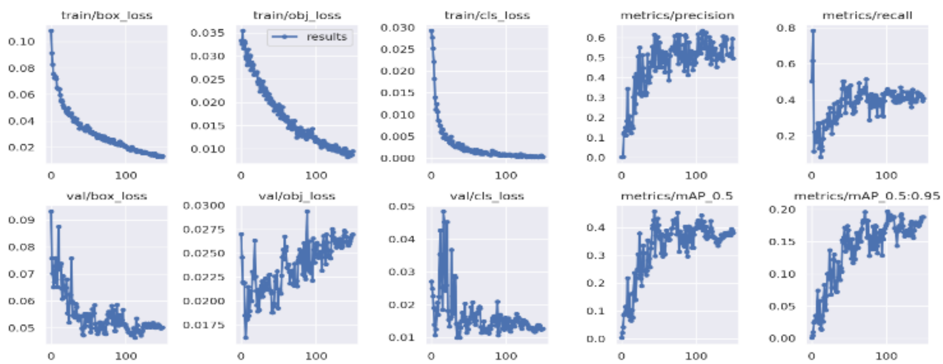


Fig. 13: Summary of metrics that occurred during image training

Figure 13 is composed of two rows of records where the first row graphically represents the training and the second row represents the validation of the indicators (box_loss, obj_loss, cls_loss), as well as the values of the metrics of the precision and the recall which is made up of a set of graphs assigned as the Box_loss.- which represents the loss that occur in the regression of the error mean square, while the Obj_loss represents values of the trust where it indicates the presence of detected objects, Cls_loss shows the loss of the classification of the classes, the graph where the Precision is shown, this indicates the measurement of the predictions of the box if they are correct, that is to say that the TP / TP + FP and finally the Recall, which measures how many true there are in box_loss, of which they were correctly predicted that is to say TP / TP + FN representing a value of mAP_0.5 that comes to be the average precision (mAP) that transits at the threshold of the intersection of the values from 0.5 to 0.95 being the average in different thresholds that exist.

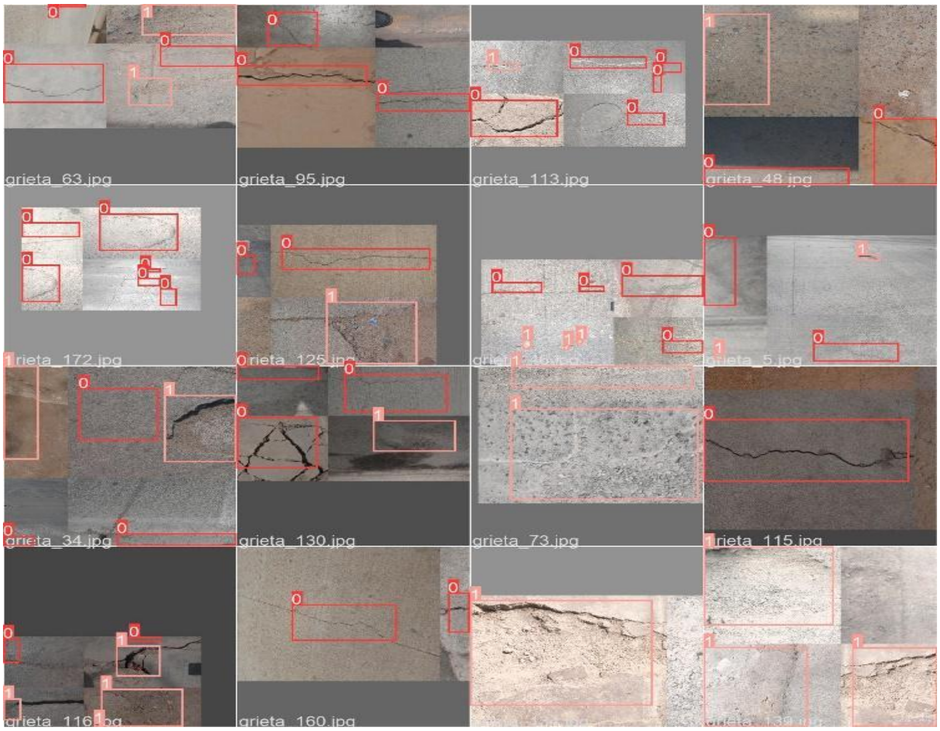


Fig. 14: Precision of cracks and gaps obtained during the training of the dataset_clasificacionGH

Figure 14 shows the image matrix, has indicators between 0 and 1 that presents a type of class detailed in the confusion matrix (0 crack - 1 hole). Likewise, as a result of the training, 36.67% were obtained that identify gaps while 63.33% identified cracks in the whole of the datase composed of images of asphalted pavements and concretes.

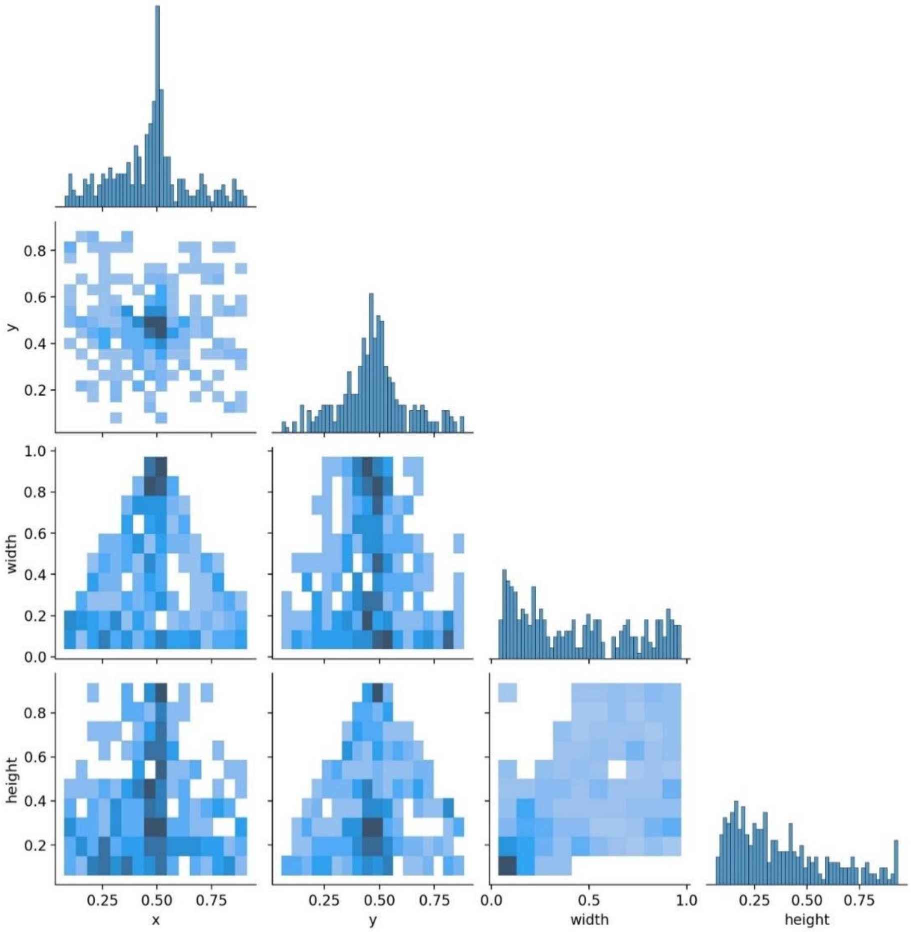


Fig. 15: Scatter matrix graph showing histograms.

Figure 15 discusses the transformation of images that occurs in the transition of the width and height resizing that processes the convolutional algorithms for each pixel that passes on the x,y axes into a value of 0.25 to pass the information to the next convolutional layer. Under this concept, the indicators that are on the x-axis of the dispersion matrix are composed of indicators from 0.25 to 0.75, while the indicators of the y axis are in values from 0.2 to 0.8.

4.2. Test Results

After the training of the dataset, the model called dataset_clasificadorGH was obtained, this model was implemented in the development of an application to detect cracks and gaps on road structures in asphalt or concrete; for which photographs of a main avenue were taken through a smartphone obtaining the following results.

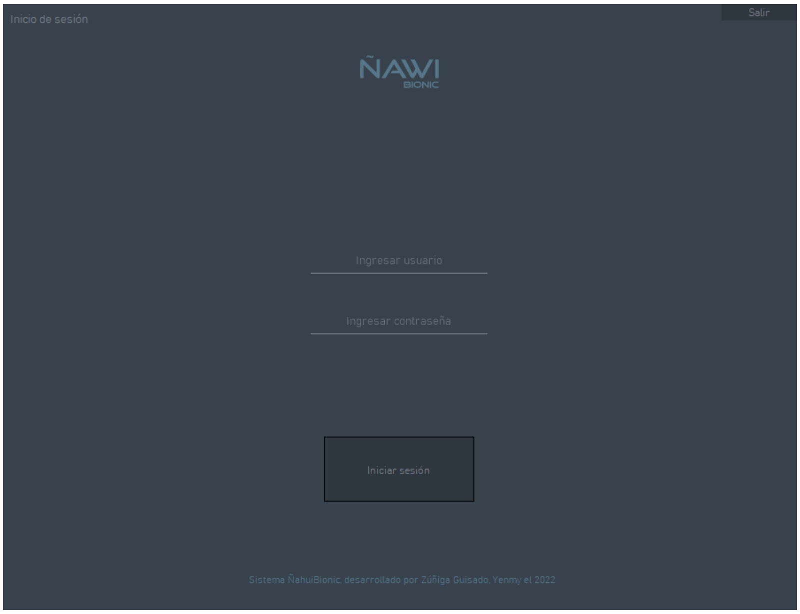


Fig. 16: User login interface

Figure 16 shows a window where you can validate the registration of the user and password for the login, developed under the Python language and unstructured Json data.

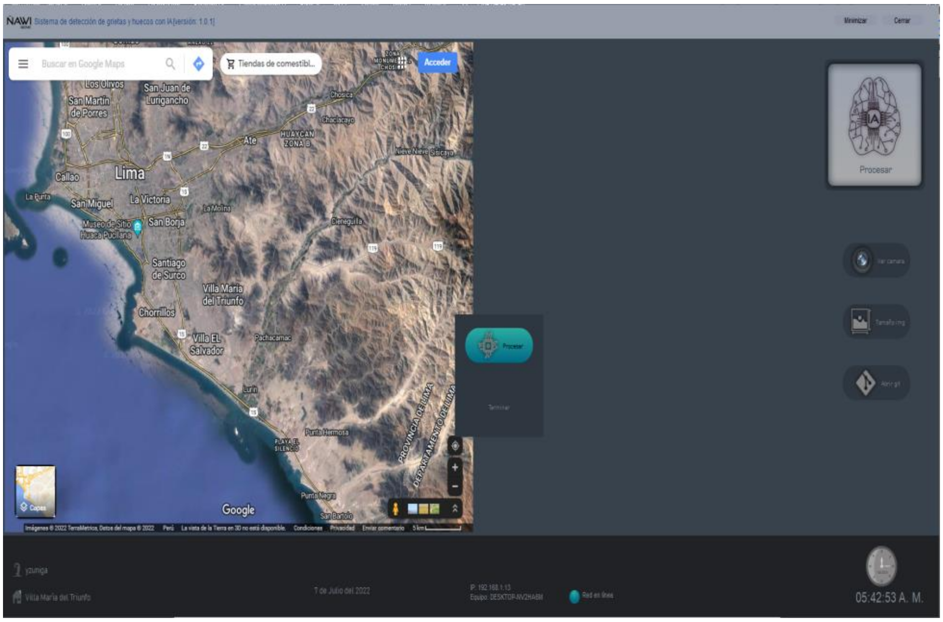


Fig. 17: Main screen to start the new model of the dataset_clasificadorGH.pt

Likewise, the main module can be seen in Figure 17, with the presence of a function that interacts with the new model of the dataset, allowing to detect cracks and gaps through the use of a camera in real time. In addition to this, this module contains functions such as capturing, saving, deleting, resizing images that allow you to manage the process, as well as the implementation of file routing through Git software and the use of the Google Maps API that allows you to geographically locate the study areas.

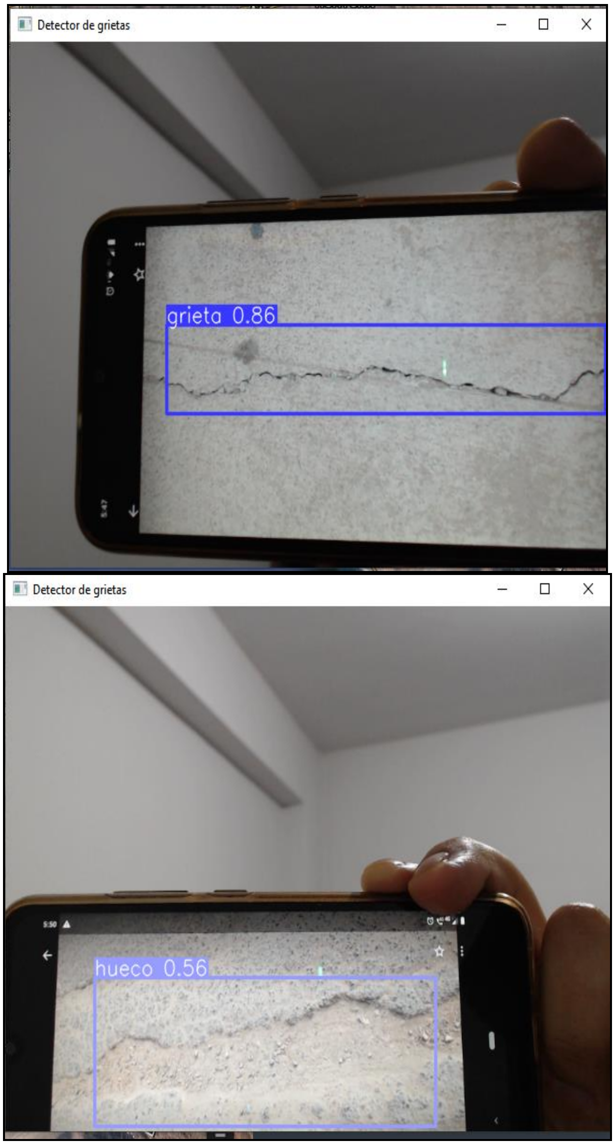


Fig. 18: Class detector window

In Figure 18, the module or window entitled "Crack Detector", this shows the results obtained from the crack accuracy at a value of 0.86, while the accuracy in gap detection was obtained at 0.56. Likewise, results of the sample of images can be obtained through smartphones, printed images and video of pavements that are composed of asphalt or concrete.

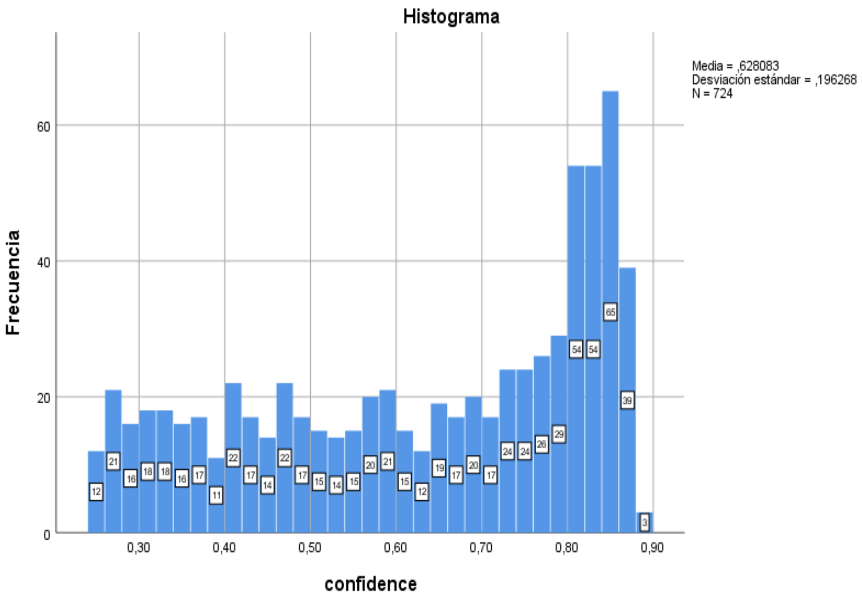


Fig. 20: Histogram of confidence over 30 test images

In Figure 20 it can be seen that the average exceeds 50%, being the indicator of 0.62, while the deviation reaches 0.19 in a total of 724 iterations.

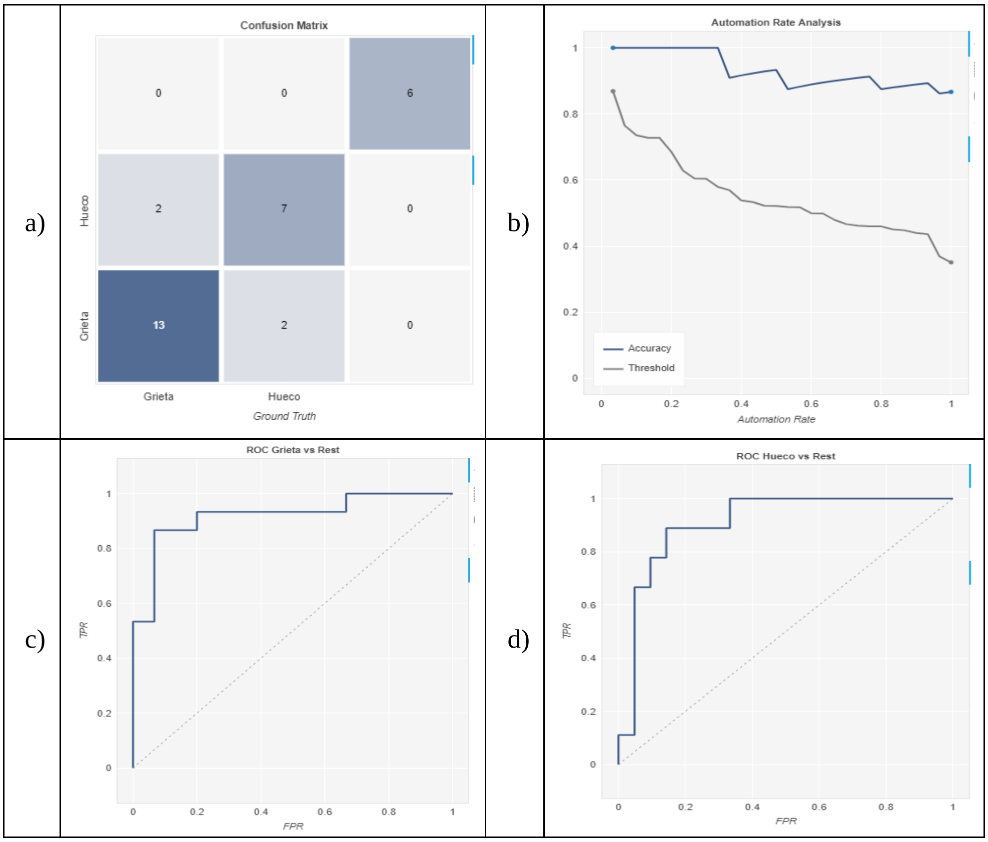


Fig. 21: Registro de métricas

In Figure 21.a, the confusion matrix is observed, where it can be seen that, of a total of 15 cracks, 13 and 2 incorrect cracks have been detected correctly; in relation to the gaps of a total of 9, 7 and 2 incorrect ones have been detected correctly; likewise, it can be seen that 6 characteristics were detected that do not correspond to hollow or crack. In Figure 21.b, the class accuracy (crack - hole) can be seen with values between 0.8 to 1. In addition, Figure 21.c and 21.d shows the ROC curve of the two classes, where a threshold higher than 0.5 is shown, being acceptable.

Table 4: Summary of Test indicators.

Metric	Type of fault	Result
Precision	Crack	0.93
	Hollow	0.77
Sensitivity	Crack	0.92
	Hollow	0.91
F1-Score	Crack	0.86
	Hollow	0.77

Table 4 shows a summary of the indicators evaluated in the test or test of the program, the averages of 85% accuracy, 91.5 sensitivity and 81.5% F1-Score were obtained.

5. Discussion

The results obtained are optimal in the detection of cracks and gaps in the pavement of road structures such as asphalt or concrete for the study area (mainly damaged areas), so a series of processes are carried out applying deep learning techniques and technical tools. As you can understand, some methods tend to study crack detection by capturing static images. In the results obtained by Fan et al. (2020), the accuracy reaches between 0.19 and 0.92, for the recall it reaches from 0.67 to 0.93, and in F1 it reaches from 0.28 to 0.92 in images with resolutions of 768x512, 991x462 and 311x462 pixels. While the metrics obtained in the work are within the reliability range of metrics between 0.32 and 0.85 in images with resolution of 1920x1080 pixels.

In the work presented by Fan et al. (2020), a dataset preparation technique is used, involving the cleaning of images containing crack and gap characteristics, as well as other types of failures that can occur on the surface of road structures to which the technique is applied. Through convolution operations based on the processing of pixel information in an image, which is crucial for deep learning, it works similarly to the techniques used to detect cracks and holes in road structures (Opara et al. 2020), being the techniques applied in the development of the present work.

The detection of cracks and holes was determined by deep learning and the use of smartphones for image taking, which follows the applied line of this research and that of Maeda et al., (2018), obtaining the objectives set favorably.

6. Conclusions

From the works applied with artificial intelligence methodologies through deep learning, the present study has positive results, applied through the model dataset_clasificadorGH in the intelligent application, allowing the detection of cracks and gaps, which appear on road traffic surfaces, such as asphalt or concrete structure pavements in different textures or states. Likewise, the model is an alternative solution to the problems that society is going through, detecting dangerous areas for traffic and that can subsequently generate statistical reports, for an adequate analysis by the competent authorities and execute an investment for the rehabilitation of these structures.

It is suggested to use different intelligent models that can be developed with convolutional neural network structure technologies and sophisticated algorithms such as YOLOv5x. In addition, in the present study indicators were obtained that exceed 50% in the accuracy of detection of cracks and gaps reaching indicators of

86%. On the other hand, during the training indicators of up to 93% accuracy were obtained which guarantees the reliability of the model for its implementation in the detection of cracks and gaps since the studies evaluated and implemented in other countries include records with indicators similar to those obtained.

As limitations of the study is the lack of a suitable device to collect the images, only a smartphone was available; in addition to the materials that allow better lighting when capturing the images with the smartphone.

References

Abeliuk, A. & Gutiérrez, C. (2021). Historia y evolución de la inteligencia artificial. 21, 14–21. Retrieved from https://scholar.google.com/scholar?hl=es&as_sdt=0%2C5&q=Historia+y+evoluci%C3%B3n+de+la+inteligencia+artificial&btnG=

Arbaoui, A., Ouahabi, A., Jacques, S., & Hamiane, M. (2021). Concrete cracks detection and monitoring using deep learning-based multiresolution analysis. *Electronics 2021*, 10(15), 1772. DOI:10.3390/ELECTRONICS10151772

Augustauskas, R. & Lipnickas, A. (2020). Improved pixel-level pavement-defect segmentationusing a deep autoencoder. *Sensors*, 20(9), 2557. DOI:https://www.mdpi.com/1424-8220/20/9/2557/htm

Backman, K., Kulic, D., & Chung, H. (2021). Learning to assist drone landings. *IEEE Robotics and Automation Letters*, 6(2), 3192–3199. DOI:10.1109/LRA.2021.3062572

Cass, S. (2021). Top programming languages: Our eighth annual probe into what's hot and not. *IEEE Spectrum*, 58(10), 17. DOI:10.1109/MSPEC.2021.9563957

Chun, C. & Ryu, S.-K. (2019). Road surface damage detection using fully convolutional neural networks and semi-supervised learning. *Sensors*, 19(24), 5501. DOI:10.3390/S19245501

Coluccia, A., Fascista, A., Schumann, A., Sommer, L., Dimou, A., Zarpalas, D., & Rajashekar, S. (2021). Drone vs. bird detection: Deep learning algorithms and results from a grand challenge. *Sensors*, 21(8), 2824. DOI:10.3390/S21082824

Enzo, R. A., Cassandra, C. R., Felipe, A. N., & Pablo, M. Q. (Eds.). (2021). Artificial intelligence: Development of classification and segmentation algorithms in chest radiography, 27. DOI:10.4067/S0717-93082021000100008

Eslami, E. & Yun, H. B. (2021). Attention-based multi-scale convolutional neural network (A+MCNN) for multi-class classification in road images. *Sensors*, 21(15), 5137. DOI:10.3390/S21155137

Fan, Z. L., Li, H., Su, C., Bruno, J., & Giuseppe, S. L. (2022). Use of parallel resnet for high-performance pavement crack detection and measurement. *Sustainability*, 14(3), 1825. DOI:https://www.mdpi.com/2071-1050/14/3/1825

Ghaleb, M. S., Ebied, H. M., Shedeed, H. A., & Tolba, M. F. (2022). Image Retrieval Based on Deep Learning. *Journal of System and Management Sciences*, 12(2), 483–502. DOI:10.33168/JSMS.2022.0226

He, Z. & Liu, Q. (2020). Deep regression neural network for industrial surface defect detection. *IEEE Access*, 8, 35583–35591. DOI:10.1109/ACCESS.2020.2975030

Instituto Nacional de Estadística e Infrmática. (2022). Indicador de la Actividad Productiva Departamental - Cuarto Trimestre 2021. Retrieved from https://m.inei.gob.pe/biblioteca-virtual/boletines/produccion-nacional-departamenal-9836/1/#lista

Janani, L., Doley, R., Sunitha, V., & Mathew, S. (2022). Precision enhancement of smartphone sensor-based pavement roughness estimation by standardizing host vehicle speed. *Canadian Journal of Civil Engineering*, 49(6), 716–730. DOI:10.1139/CJCE-2021-0182

Jansen, P. (2022). TIOBE - the software quality company. Retrieved October 9, 2021, from tiobe.com website: https://www.tiobe.com/tiobe-index/

Kim, Y., Song, H., & Han, J. (2022). A deepfake-based deep learning algorithm for medical data manipulation detection. *Journal of System and Management Sciences*, *12*(1), 13–26. DOI:10.33168/JSMS.2022.0102

Kittipongdaja, P. & Siriborvornratanakul, T. (2022). Automatic kidney segmentation using 2.5D ResUNet and 2.5D DenseUNet for malignant potential analysis in complex renal cyst based on CT images. *EURASIP Journal on Image and Video Processing*, 15. DOI:10.1186/s13640-022-00581-x

Kumar, P., Batchu, S., Swamy S., N., & Kota, S. R. (2021). Real-time concrete damage detection using deep learning for high rise structures. *IEEE Access*, 9, 112312–112331. DOI:10.1109/ACCESS.2021.3102647

López, M. T., Fernández-Caballero, A., Mira, J., Delgado, A. E., & Fernández, M. A. (2006). Algorithmic lateral inhibition method in dynamic and selective visual attention task: Application to moving objects detection and labelling. *Expert Systems with Applications*, 31(3), 570–594. DOI:10.1016/j.eswa.2005.09.062

Madasamy, K., Shanmuganathan, V., Kandasamy, V., Lee, M. Y., & Thangadurai, M. (2021). OSDDY: Embedded system-based object surveillance detection system with small drone using deep YOLO. *EURASIP Journal on Image and Video Processing*, 1(1), 1–14. DOI:10.1186/S13640-021-00559-1

Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018). Road damage detection using deep neural networks with images captured through a smartphone. University of Tokyo, 14

Morales, T., Sarabakha, A., & Kayacan, E. (Eds.). (2020). image generation for efficient neural network training in autonomous drone racing. DOI:10.1109/IJCNN48605.2020.9206943

Mukhopadhyay, S., Cennamo, N., Deen, M. J., Lee, J., Fabio, T. S., Chen, C., & Seo, H. (2021). Deep learning-based thermal image analysis for pavement defect detection and classification considering complex pavement conditions. *Remote Sensing*, 14(1), 106. DOI:10.3390/RS14010106

Oliphant, T. (2022). Numpy. Retrieved March 20, 2022, from numpy.org website: https://numpy.org/doc/stable/index.html#

Opara, J. N., Thein, A. B. B., Izumi, S., Yasuhara, H., & Chun, P.-J. (2021). Defect detection on asphalt pavement by deep learning. *International Journal of Geomate*, 21(83), 87–94. DOI:10.21660/2021.83.6153

Palma, J. T., & Morales, R. M. (2008). Inteligencia artificial. técnicas, métodos y aplicaciones. McGRAW-HILL/INTERAMERICANA DE ESPAÑA, S. A. U

Qiao, W., Liu, Q., Wu, X., Ma, B., & Li, G. (2021). automatic pixel-level pavement crack recognition using a deep feature aggregation segmentation network with a scSE attention mechanism module. *Sensors*, 21(9), 2902. DOI:10.3390/S21092902

Qt Group. (2022). Qt. Retrieved from Qt The Qt Company website: https://investors.qt.io/?hsLang=en

Ramalingam, B., Hayat, A. A., Elara, M. R., Gómez, B. F., Yi, L., Pathmakumar, T., & Subramanian, S. (2021). Deep learning based pavement inspection using self-reconfigurable robot. *Sensors*, 21(8), 2595. DOI:10.3390/S21082595

Sarmiento-Ramos, J. L. (2020). Aplicaciones de las redes neuronales y el deep learning a la ingeniería biomédica. Revista UIS Ingenierías, 19(4), 1–18. DOI:10.18273/REVUIN.V19N4-2020001

Shatnawi, N. (2018). Automatic pavement cracks detection using image processing techniques and neural network. *International Journal of Advanced Computer Science and Applications*, 9(9), 399–402. DOI:10.14569/IJACSA.2018.090950

Shim, S., Kim, J., Cho, G. C., & Lee, S. W. (2020a). Multiscale and adversarial learning-based semi-supervised semantic segmentation approach for crack detection in concrete structures. *IEEE Access*, 8, 170939–170950. DOI:10.1109/ACCESS.2020.3022786

Silva, L. A., Blas, H. S. S., García, D. P., Mendes, A. S., & González, G. V. (2020). An architectural multi-agent system for a pavement monitoring system with pothole recognition in UAV images. *Sensors*, 20(21), 6205. DOI:10.3390/S20216205

Skalski, P. (2022). Make alpha sense. Retrieved June 2, 2022, from Make Alpha Sense.ai website: https://www.makesense.ai/

Srivastava, S., Divekar, A. V., Anilkumar, C., Naik, I., Kulkarni, V., & Pattabiraman, V. (2021). Comparative analysis of deep learning image detection algorithms. *Journal of Big Data*, 8(1), 1–27. DOI:10.1186/S40537-021-00434-W

Sun, Y., Yang, Y., Yao, G., Wei, F., & Wong, M. (2021). Autonomous crack and bughole detection for concrete surface image based on deep learning. *IEEE Access*, 9, 85709–85720. DOI:10.1109/ACCESS.2021.3088292

Van Rossum, G. (2022). Python. Retrieved September 7, 2021, from www.python.org website: https://www.python.org/

Xiao, Y., Daniel, L., & Gashinova, M. (2021). Image segmentation and region classificationin automotive high-resolution radar imagery. *IEEE Sensors Journal*, 21(5), 6698–6711. DOI:10.1109/JSEN.2020.3043586

Zhang, L. (2021). Hand-drawn sketch recognition with a double-channel convolutional neural network. *EURASIP Journal on Advances in Signal Processing*, 2021(1), 1–12. DOI:10.1186/S13634-021-00752-4

Zhang, M., Liu, Y., Luo, S., & Gao, S. (2020). Research on Baidu street view road crack information extraction based on deep learning method. *Journal of Physics: Conference Series*, 1616(1), 012086. DOI:10.1088/1742-6596/1616/1/012086