GPR image interpretation advancement for smarter technical management of water leakage in urban water infrastructures

Ayala-Cabrera, David, Izquierdo, Joaquín

David Ayala-Cabrera, Joaquín Izquierdo, "GPR image interpretation advancement for smarter technical management of water leakage in urban water infrastructures," Proc. SPIE 11863, Earth Resources and Environmental Remote Sensing/GIS Applications XII, 118630S (12 September 2021); doi: 10.1117/12.2599943

SPIE.

Event: SPIE Remote Sensing, 2021, Online Only

GPR image interpretation advancement for smarter technical management of water leakage in urban water infrastructures

David Ayala-Cabrera*[a], Joaquín Izquierdo[b]
[a]CWRR-School of Civil Engineering, University College Dublin, Richview Newstead Belfield Dublin 4, Dublin, Ireland; [b]FluIng-IMM, Universitat Politècnica de València, Camino de Vera s/n, Valencia, Spain.

ABSTRACT

This paper combines such powerful non-destructive technique as ground penetrating radar (GPR) with intelligent data analysis in order to acquire new knowledge on mapping/monitoring/verification systems aimed at initially facilitating the evaluation of the health of the buried assets of water distribution system (WDS) infrastructures in urban areas. Pre-processing techniques based on a multi-agent approach coupled with a suitable analysis of the properties of the obtained groups of objects found, as well as a classification supported by machine learning techniques, are presented in this work. The work is based on GPR image studies conducted under controlled laboratory conditions using a commercial antenna. The buried objects correspond to pipes of various materials commonly used in WDSs. Furthermore, GPR images of various pipes either empty or full of water, as well as leaking water, are also included in the dataset of the study. The dataset is divided into two subsets that pursue: 1) the development of the methodology; *i.e.* capturing objects in GPR images to favour the feature extraction process, and 2) the evaluation of the feasibility of implementation of the proposed classification, as well as the response of the methodology to various environmental interactions recorded in the images. The results of this work are promising in the sense of promoting the inclusion of powerful tools such as GPR towards the provision of smarter tools that adequately support technical management in WDSs, and, eventually, that of other surrounding infrastructure(s).

Keywords: Ground penetrating radar, semi-autonomous/autonomous GPR image interpretation, intelligent data analysis, water distribution systems, water leakage.

1. INTRODUCTION

Water leaks are a major concern in water distribution systems (WDSs). This comes from the fact that appropriate control/monitoring of those undesired events can ensure sustainability and efficient use of water resources in the system. Accordingly, there is an urgent need to improve effective detection of water leaks in WDSs. Three aspects are essential in the detection of water leaks in the urban infrastructure of a WDS: 1) the detection of the leak itself, 2) the assessment of the pipe leaking, and 3) the evaluation of the surrounding infrastructure(s). These three aspects are essential for an adequate knowledge of the urban infrastructure assets that allows actions to be undertaken in order to remedy water losses. This is possible because non-destructive methods are powerful tools that may provide WDSs with reliable and affordable information about the health of their assets.

The ground penetrating radar (GPR) is of particular interest for this work due to its demonstrated capacity for effective data collection of buried assets not only in water infrastructures but also in the health evaluation of other related infrastructures *e.g.* transport infrastructures[1],[2]. This derives from the versatility of the GPR for subsoil inspection, and identification/monitoring of buried assets with varied characteristics, compared to other non-destructive methods (for example, identification of plastic pipes in WDSs). The use of GPR for leak detection is well established in the literature (*e.g.* De Coster et al.[3]; Amran et al.[4]; Ocaña-Levario et al.[5]; Ayala-Cabrera et al.[6]), and is motivated by its characteristics as a non-destructive geophysical method that allows to evaluate both non-metallic and metallic objects and soil materials.

*david.ayala-cabrera@ucd.ie; phone +353 1 716 3208; https://people.ucd.ie/david.ayala-cabrera

Earth Resources and Environmental Remote Sensing/GIS Applications XII,
edited by Karsten Schulz, Proc. of SPIE Vol. 11863, 118630S · © 2021 SPIE
CCC code: 0277-786X/21/$21 · doi: 10.1117/12.2599943

Hyperbolas are the most common shapes obtained in GPR images from subsurface surveying[7]. Consequently, hyperbola characterization is among the most common object identification strategies in GPR studies. However, the first reflector eventually provides limited information about the target object[8]. This is due to the fact that sometimes more than one pair of reflections may appear for the same initial reflector, from which multiple interpretations for the same image may be obtained[9]. The interpretation of the GPR images is commonly conducted by expert personnel, thus limiting its potential for use. The difficulty in interpreting GPR images, added to human error, and the huge amount of information generated hinder the extensibility of the use of this type of tools. In this context, an appropriate interpretation of the data obtained by non-destructive methods seems to be an essential requirement. This work tackles, as a key factor, some interpretability enhancement of GPR images to evaluate WDS assets. The aim is to advance towards the generation of semi-autonomous/autonomous identification and characterization tools ready to be used by non-expert operators in GPR image interpretation. To this end, the methodology proposed in this paper analyses images from subsoil surveys obtained by using GPR methods, and generates a classification of the objects embedded into the GPR images. The proposed classification uses pre-processing techniques based on a multi-agent-based approach and is complemented by the analysis of the properties of the obtained groups (*e.g.* area, centroid, perimeter). Afterwards, a classification of the selected property via machine learning techniques is carried out. The work is based on GPR image studies conducted under controlled laboratory conditions using a commercial antenna. The buried objects correspond to pipes of various materials commonly used in WDSs. Furthermore, GPR images from various pipes, either empty or full of water, as well as leaking water, are also included in the dataset of the study. The dataset is divided into two subsets that pursue: 1) the development of the methodology; *i.e.* capturing objects in GPR images to favour the feature extraction process, and 2) the evaluation of the feasibility of implementation of the proposed classification; as well as the response of the methodology to various environmental interactions recorded in the images.

2. METHODS

Let us understand the GPR radargram (A) as a set Af formed by the triple $\{g, fl, r\}$ of groups (families) as in Ayala-Cabrera[10]. Where g corresponds to the families of objects embedded into the images, fl to the medium and r represents the noise. In GPR object detection, fl and r are commonly predominant families that make successfully identification of g families difficult. To cope with this situation, this work is focused on removal of both families (*i.e.* fl and r), while the g families are preserved in the final output. The removal of these families supports further processes for successful detection of GPR objects by means of reducing the dimensionality of the problem, as we can see throughout the document. This analysis approach was first proposed in Ayala-Cabrera et al.[11] and is improved here with the inclusion of new data for the training of the classifier, and their results are also extended for the water leakage problem.

2.1 GPR image pre-processing

Remotion of straight horizontal lines – Agent race pre-process. The A matrix (size of $m \times n$) that represents the initial raw GPR image in this paper, is a stack of sequential n traces captured (represented as m-vectors, X_j with $j = 1, \dots, n$) in depth which contains the electromagnetic properties of the underground prospected: $A = [X_1, \dots, X_n]$. The analysis of each n trace may be conducted either in a vertical (*i.e.* through the columns) or in a horizontal manner (*i.e.* trough the rows) of matrix A. The selection of the direction of the analysis can provide different perspectives of the problem that can be useful for its solution.

Breaking the space. To break the image space in order to form the families, the algorithm agent race[12] was used and the analysis was conducted horizontally. The output of this algorithm, which is an array ($m1 \times n$; with $m1$ the total movements for the competition) that contains either the vertical position of the peaks (output 1) or the wave amplitude value (output 2) for each agent movements is used herein to build the matrix, of size ($m \times n$), that we denoted as A_{pre} (when the competition was developed in the columns as parallel tracks to run).

The rules applied to build A_{pre} are: for positions from output 1 (for each column) whose respective values in output 2 follow the relationships: $\text{output } 2_{\text{previous}} < \text{output } 2_{\text{current}}$ are labelled as 1, if $\text{output } 2_{\text{previous}} > \text{output } 2_{\text{current}}$ as -1, and otherwise as 0 in A_{pre}. The "previous" and "current" subscripts represent the position in each column of output (output 1 or output 2), and lies from 1 to $m1 - 1$ and from 2 to $m1$; respectively. In such a manner, A_{pre} presents each trace characterised when the monotonically tendency is increasing as 1, is decreasing as -1 and, for the proper peaks and errors (*e.g.* clipped waves) as 0.

As mentioned above, the analysis in this paper was developed in a horizontal manner. To do this, before applying the agent race algorithm, matrix A is transposed, so that the described rules to build A_{pre} are applied to the new columns. After the agent competition, A_{pre} is transposed again with which the matrix A_{prerow} is obtained. In this matrix the agents perform the competition through rows, as parallel tracks, instead of through columns, as usual.

GPR pre-processing – Generation of families.

In order to generate the families for two different points of view, two matrices (both with size $m \times n$) are generated from A_{prerow} that contain the monotonic tendency to increase (denoted as U) and decrease (denoted as D) of the wave amplitude value with respect to the traces of the neighbors. Each element of U and D matrices is labeled as 1, if its corresponding value in A_{prerow} is 1 or 0 for U matrix, and -1 or 0 for D matrix. In both matrices the label 0 for each element corresponds to the remaining value of A_{prerow} (*i.e.* A_{prerow} = -1 for U, and $A_{prerow} = 1$ for D).

For families' extraction from U and D matrices the Matlab's `bwboundaries` function was used. In-depth details of the algorithm Moore-Neighbour tracing algorithm modified with the Jacob's stopping criteria that supports this function can be find in Gonzales et al.[13]. It is possible to obtain the properties of the groups found (*e.g.* area, centroid, perimeter, eccentricity, among others) in a vector V, *e.g.* $V(centroid(Xc, Yc)$, and it is in the analysis of these properties for each group that the GPR image cleaning process (classification and removal of the families g) is based.

Dataset and data labelling. The dataset used in this document for the training purposes consists of GPR tests of the pipes commonly used in WDSs (asbestos cement, cast iron, polyethylene and polyvinyl chloride (PVC)) of different materials and at different depths (from 10.5cm to 62cm) buried in a laboratory setup. The central antenna frequency corresponds to 1.5GHz.

In this paper we use a thresholding for the value for the area property of the groups in order to efficiently label some families of objects as 1 and others as 0: specifically, when the value of the area of the group obtained is below the threshold, the group is labelled as 0, and 1 otherwise. The threshold used corresponds to the value of 2568 (samples × trace) and was obtained by the average values for the groups formed by the GPR test without embedded objects (56308 groups in total). A total of 424271 groups were obtained with these tests, of which 1259 were labelled as object families and used for the classification. In addition, it should be mentioned that the number of traces obtained varies from 214 to 294 samples/image. However, additional criteria to refine the initial classification for the training of the removal process (and eventually for further classification) in order to extract and/or characterise the families of objects could be the uses or contrast images as in Ayala-Cabrera et al.[14] or Ocaña-Levario et al.[5] or even manual classifications as previously conducted.

Based in the observations Ayala-Cabrera et al.[11] proposed a change of space for the Xc component of the group property of the centroid.

This variant in the variable proposes the generation of a vector denoted as $Xmul$ that is constructed from the values of Xc. Based in the positions for Xc for this property in V, and with the help of the agent race algorithm a comparison vector with Xc is constructed, that is, $\{Xc - Xmul\}$. This variation not only facilitates the extraction of the patterns but also increases their quality by considering the peaks obtained in the data series; and, as a result, Xc goes from being a value to being a value compared to a trend.

2.2 GPR image cleaning process – Features extraction

For the cleaning process and in order to extract the families of objects, in this section we use a classification performed by a perceptron neural network. The perceptron is the simplest structure for neural networks in pattern classification problems[15].

The formulation for the perceptron in this paper (Equation 1) takes the initial values of the properties of the groups obtained (or a pre-processed values of them *i.e.* $\{Xc - Xmul\}$), as a set of predictors (in a vector form) V_i, with $i = 1, \dots, npro$, in order to provide the prediction in a vector of classes ($Clas_{PR}$).

The total number of predictors ($npro$) used may vary according to the complexity of the patterns into the predictors used, in this case we have used as a predictor the change of the space proposed in Section 2.1., for Xc, *i.e.* $\{Xc - Xmul\}$; w_0 y w_i represents the bias and weights obtained in the train process in order to make the prediction, respectively.

$$Clas_{PR} = w_0 + \sum_{i=1}^{npro} V_i w_i \ , \tag{1}$$

The mean square error (MSE) function (Equation 2) is used to compare the classes obtained in the prediction ($Clas_{PR}$) and the expected classes ($Clas_{EX}$) for each group ($k = 1, \dots, np$) for all the groups used (np). This function allows to obtain a unique value ($error$) in order to estimate the success in the prediction performed (quality of prediction).

$$error \equiv \sum_{k=1}^{np} \left[e^{(k)}\right]^2 = \sum_{k=1}^{np} \left[Clas_{PR}^{(k)} - Clas_{EX}^{(k)}\right]^2 \ . \tag{2}$$

3. CASE STUDY – WATER LEAKAGE EXAMPLE

In this section we have used the data from a simulated leak presented in Ayala-Cabrera et al.[6] in order to compare the results of the prediction. The database consists of captures of GPR images of a plastic pipe buried in a laboratory tank (Figure 1) while leaking water. One of the analysis conducted in that work consisted of the image-by-image contrast between the configuration when the system is not leaking (initial state) and leaking (final state), *i.e.* contrast method. Based on the results of the contrasts obtained, this section evaluates the effectiveness of the cleaning process described above in water leakage.

It should be mentioned that this dataset has not been used in the training of the neural network perceptron. Furthermore, the parameters of the equipment differ from the data used for training. Here we have used a parameter setting of 512samples/trace and 20ηs/512 samples, instead of 2048 samples / trace and 10ηs/2048samples for the training.

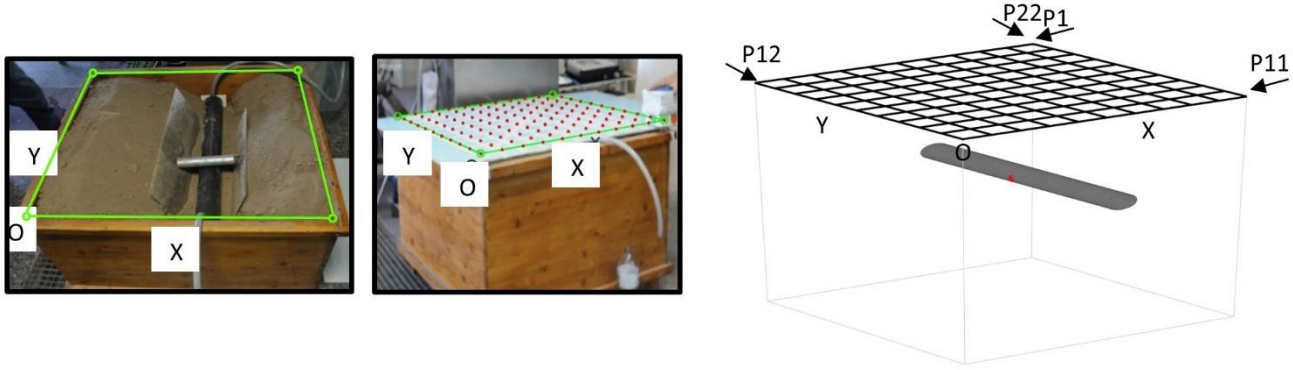

Figure 1. Setup for water leak capture using GPR in a laboratory tank.

Figure 2 and Figure 4 present the results of the image cleaning process, for profiles from P1 to P11 and from P12 to P22; respectively. Figure 3 and Figure 5 shows the results of the contrast method for the profiles between P1 to P11 and P12 to P22; respectively. In the images presented in Figure 2 and Figure 4, a resizing of each image in distance (*i.e.* abscissa axis) was not carried out. This is in order to evaluate the effectiveness of the method described in this paper on images without any previous treatment. Nevertheless, on the abscissa axes for each image an orientation distance is presented based on the identification of the tank walls for the initial and final abscissae (0.0m and 1.0m, respectively) for both Figure 2 and Figure 4. Similarly, an intermediate point has been set for the images in both Figure 2 (based on the centre of the pipe *i.e.* 0.6m) and Figure 4 (based on the crossing of the reflections of the tank walls *i.e.* 0.5m).

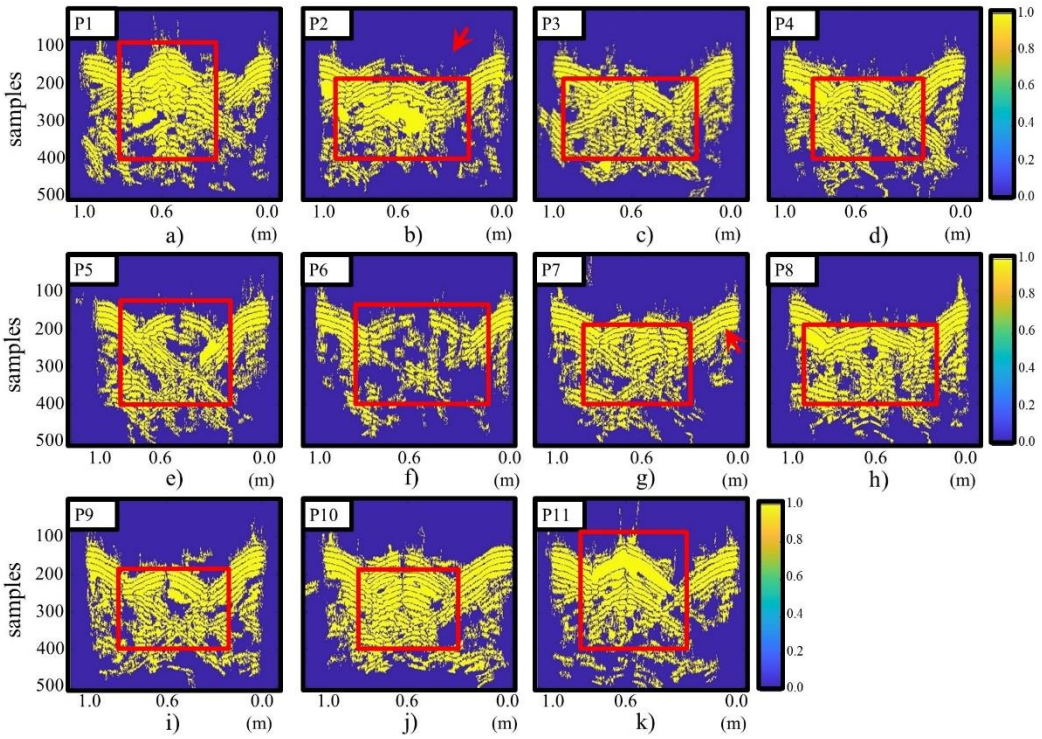

Figure 2. Image cleaning process. (a-k) Profiles P1 to P11 of the classes predicted using perceptron neural network for GPR images when the system is leaking.

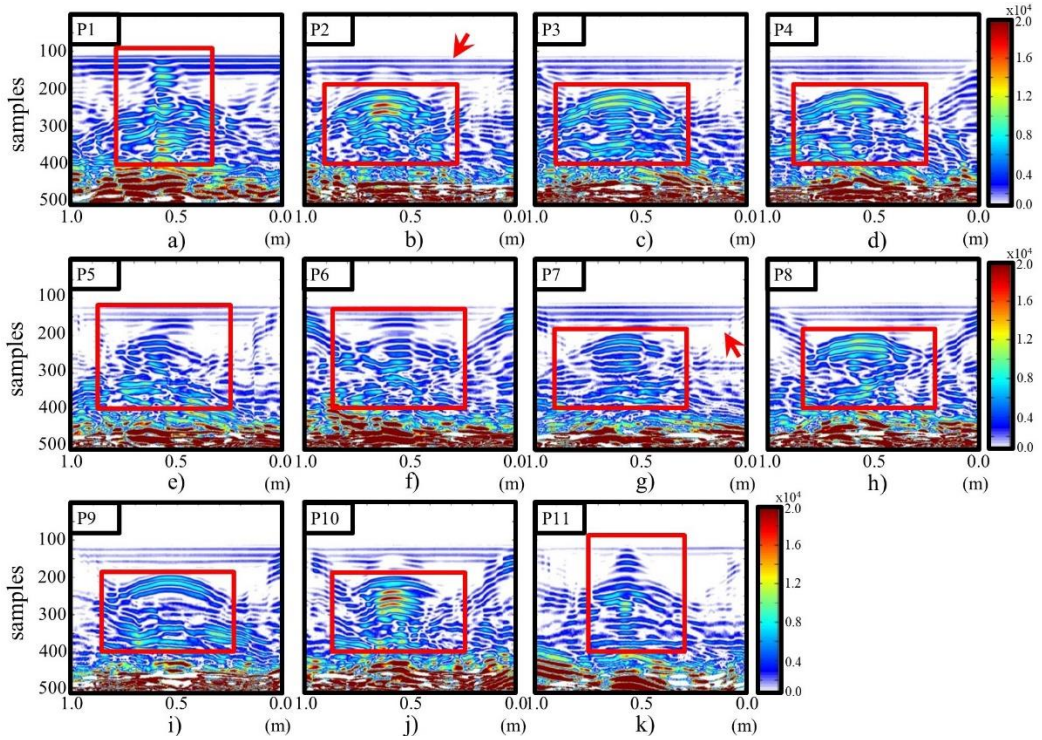

Figure 3. Contrast method. (a-k) Profiles P1 to P11 of the contrast for GPR images when the system is leaking and not leaking. Adapted from Ayala-Cabrera et al.[6].

After the application of the cleaning process described in this work, Figure 2 and Figure 4 show a successful tendency to remove both the families fl (in particular straight horizontal lines) and the families r (noise) from the initial GPR image. An example of these families' removal can be observed by comparing the red arrow located in Figure 2(b) and Figure 3(b). In Figure 3(b), despite the image is the result of a comparison between states of the system (initial and final state of the system), the horizontal lines still persist in the result.

In the case of reflections generated by the tank walls (see examples in Figure 2g and Figure 3g, red arrows), the removal is carried out more effectively in Figure 3 and Figure 5. This is due to the existence of these same reflections both in the image without leakage and with leakage cases. However, by considering that these reflections are properties of the wooden tank wall object (even if they are not the target object), it could be considered as successful the prediction obtained. The classification in such cases, could be referred to further and more specific classification processes, *e.g.* material classification, among others. The analysis of this type of vertical objects can be of interest in WDS due to the existence in real systems of components, that need to be characterised[16] (such as manholes, electrical boxes, basement walls, among others), which can be affected by the failure of the pipe through leaks of water.

Based on the results obtained in Figure 3, we present in a red box the pipe that conducts the water for Figure 2 and Figure 3. This box shows us the components of the system (with water) such as the vertical pipe inlet (inset a), the drilled pipe (insets from b to j), and the vertical pipe outlet (inset k) for both Figures. The prediction for these objects was considerably successful (Figure 2), since that it collects a number of the characteristics of the water pipe system obtained by means of the contrast method (Figure 3). In this sense, it should be noted the variety of shapes obtained with the prediction (and even with the contrast method), which go beyond the hyperbolas on which a considerable number of studies are based for the identification of pipes or subsurface components (*e.g.* Lei et al.[17]; Feng et al.[18]). The consideration of these characteristics may lead to more information regarding the health of the water pipes.

In Figure 4e, f, g (see red box), the footprint of the water leaked from the pipe have been detected in a very similar way to that obtained by the contrast method (Figure 5e, f, g; see red box). The groups that make up the water leaks present a more defined structure, being more distinguishable than their respective results for the contrast method. Two additional groups are observed in profiles P12 and P22 which correspond to an unknown element are observed in Figure 4a and b (green box). These elements are not reflected in their respective images in the figure of the contrast method, therefore they are not attributed to the water footprint.

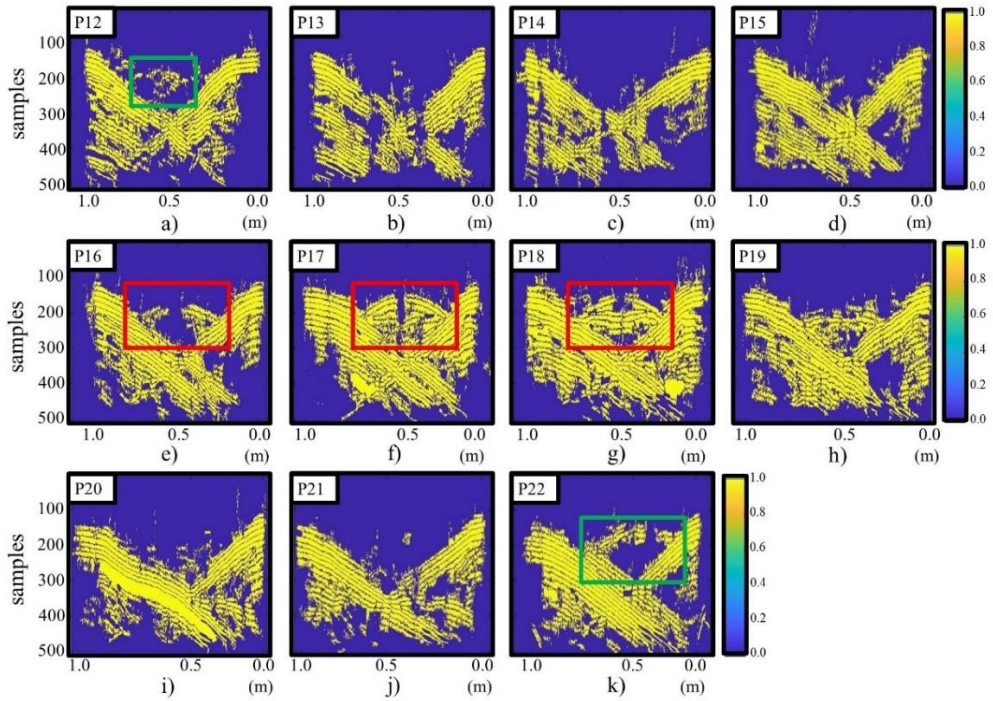

Figure 4. Image cleaning process. (a-k) Profiles P12 to P22 of the classes predicted using perceptron neural network for GPR images when the system is leaking.

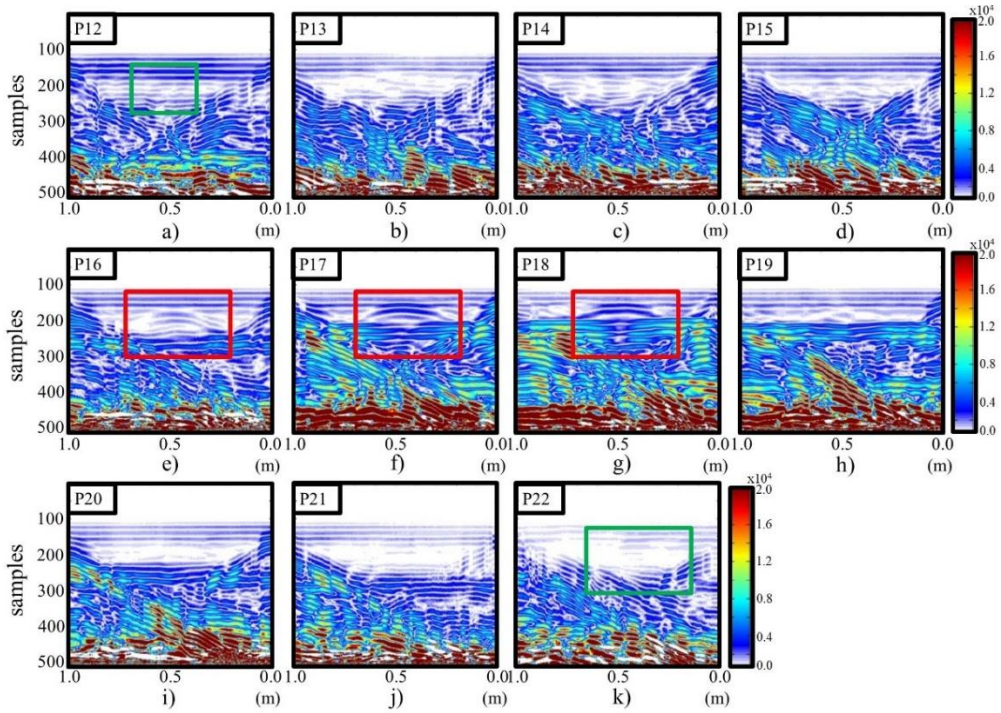

Figure 5. Contrast method. (a-k) Profiles P12 to P22 of the contrast for GPR images when the system is leaking and not leaking. Adapted from Ayala-Cabrera et al.[6].

4. CONCLUSIONS

This work presents a methodology based on multi-agent systems and classification using perceptron neural networks for the generation of the groups and their extraction from GPR images. The methodology is an easy-to-apply procedure that derives in the cleaning of the image, while the relevant characteristics of the target objects are preserved (in our case water leakage and pipe objects). The cleaning process was applied on a water leak dataset collected with a commercial GPR antenna in a laboratory setting. The results obtained for the objects extraction were also compared with a contrast method in order to validate de results. The validation showed promising results as similar results were obtained in terms of the quality of the extracted patterns for the analysed objects, without the need to use a reference image. In the same way, the patterns obtained are purer in terms of noise reduction, thus facilitating a more robust analysis process. The results of this study are promising in the sense that the classification of areas in GPR images allows the extraction of relevant characteristics of WDSs' components, what enables to advance towards semi-autonomous/autonomous interpretation of those datasets, while reducing experts' dependence in the interpretation of this information.

REFERENCES

[1] Tosti, F., Gagliardi V., D'Amico, F. and Alani A., "Transport infrastructure monitoring by data fusion of GPR and SAR imagery information," Transp Res Proc 2020, 32, 771-778 (2020).

[2] Bianchini Ciampoli, L., Gagliardi, V., Clementini, C., Latini, D., Del Frate, F., and Benedetto, A., "Transport Infrastructure Monitoring by InSAR and GPR Data Fusion," Surv Geophys 41, 371–394 (2020).

[3] De Coster, A., Pérez Medina, J.L., Nottebaere, M., Alkhalifeh, K., Neyt, X., Vanderdonckt, J., and Lambot, S., "Towards an improvement of GPR-based detection of pipes and leaks in water distribution networks," Journal of Applied Geophysics. 162, 138-151 (2019).

[4] Amran, T., Ismail, M., Ahmad, M., Amin, M., Sani, S., Masenwat, N., Ismail, M., and Hamid, S., "Detection of underground water distribution piping system and leakages using ground penetrating radar (GPR)," Proc. of the AIP Conference, vol. 1799, 030004.1–030004.10 (2017).
[5] Ocaña-Levario, S.J., Ayala-Cabrera, D., Izquierdo, J., and Pérez-García, R., "3D model evolution of leak based on GPR image interpretations," Water Science & Technology: Water Supply 15 (6), 1312-1319 (2015).
[6] Ayala-Cabrera, D., Herrera, M., Izquierdo, J., Ocaña-Levario, S.J., and Pérez-García, R., "GPR-based water leak models in water distribution systems, " Journal Sensors 2013, 15912-15936 (2013).
[7] Ahmadi, R., Fathianpour, N., and Norouzi, G.H., "Detecting physical and geometrical parameters of some common geotechnical targets through their effects on GPR responses," Arab. J. Geosci. 8(7), 4843-4854 (2015).
[8] Janning, R., Busche, A., Horváth, T., and Schmidt-Thieme, L., "Buried pipe localization using an iterative geometric clustering on GPR data," Art. Intell. Rev. 42(3), 403-425 (2014).
[9] Robinson, M., Bristow, C., McKinley, J., and Ruffell, A., Ground penetrating radar, Part 1, Sec: 5.5. In Clarke, L.E & Nield, J.M. (Eds.) Geomorphological Techniques (online Edition). British Society for Geomorphology, London (UK), pp. 1–26 (2013).
[10] Ayala-Cabrera, D., Characterization of components of water supply systems from GPR images and tools of intelligent data analysis, Ph.D. thesis, Universitat Politècnica de València, Valencia, Spain (2015).
[11] Ayala-Cabrera, D., Herrera, M., Izquierdo, J., Pérez-García, R., and Piller, O., "A new GPR image analysis proposal based on a multi-agent approach and properties of groups - Towards automatic interpretations," Proc. 8th International Congress on Environmental Modelling and Software, iEMSs 2016 (2016a).
[12] Ayala-Cabrera, D., Izquierdo, J., Montalvo, I., and Pérez-García, R., "Water supply system component evaluation from GPR radargrams using a multi-agent approach," Mathematical and Computer Modelling 57(7-8), 1927-1932 (2013).
[13] Gonzales, R.C., Woods, R.E., Eddins, S.L., "Digital image processing using MatLab", Pearson Prentice Hall, Upper Saddle River, NJ, USA (2004).
[14] Ayala-Cabrera, D., Ocaña-Levario, S.J., Izquierdo, J., and Pérez-García, R., "3D Representation of (buried) water supply elements using pre-processed GPR images," Athens Journal of Technology & Engineering 3(3), 241-254 (2016b).
[15] Datsi, T., El Oirrak, A., and Aznag K., "Classification of iris plant using perceptron neural network". Proc. Ezziyyani M. (eds) Advanced Intelligent Systems for Sustainable Development (AI2SD'2019). AI2SD 2019. Advances in Intelligent Systems and Computing, vol 1105. Springer, Cham (2020).
[16] Kim, N., Kim, K., An, Y.K., Lee, H.J., and Lee, J.J., "Deep learning-based underground object detection for urban road pavement". International Journal of Pavement Engineering 21(13), 1638-1650 (2020).
[17] Lei, W., Luo, J., Hou, F., Xu, L., Wang, R., and Jiang, X., "Underground cylindrical objects detection and diameter identification in GPR B-Scans via the CNN-LSTM Framework," Electronics 9(11), 2-16 (2020).
[18] Feng, J., Yang, L., Wang, H., Tian, Y., and Xiao, J., "Subsurface Pipes Detection Using DNN-based Back Projection on GPR Data," Proc. 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 266-275 (2021).