WARSAW UNIVERSITY OF TECHNOLOGY Index 351733
FACULTY OF CIVIL ENGINEERING
COMMITTEE FOR CIVIL AND WATER ENGINEERING
POLISH ACADEMY OF SCIENCES ISSN 1230-2945

DOI: 10.24425/ace.2024.149851

ARCHIVES OF CIVIL ENGINEERING

Vol. LXX ISSUE 2 2024

pp. 59 –77



**Research paper**

# Automatic classification of underground utilities in Urban Areas: A novel method combining ground penetrating radar and image processing

**Klaudia Pasternak[1], Anna Fryśkowska-Skibniewska[2]**

**Abstract:** Precise determination of the location of underground utility networks is crucial in the field of civil engineering for: the planning and management of space with densely urbanized areas, infrastructure modernization, during construction and building renovations. In this way, damage to underground utilities can be avoided, damage risks to neighbouring buildings can be minimized, and human and material losses can be prevented. It is important to determine not only the location but also the type of underground utility network. Information about location and network types improves the process of land use design and supports the sustainable development of urban areas, especially in the context of construction works in build-up areas and areas planned for development. The authors were inspired to conduct research on this subject by the development of a methodology for classifying network types based on images obtained in a non-invasive way using a Leica DS2000 ground penetrating radar. The authors have proposed a new classification algorithm based on the geometrical properties of hyperboles that represent underground utility networks. Another aim of the research was to automate the classification process, which may support the user in selecting the type of network in images that are sometimes highly noise-laden.

The developed algorithm shortens the time required for image interpretation and the selection of underground objects, which is particularly important for inexperienced operators. The classification results revealed that the average effectiveness of the classification of network types ranged from 42% to 70%, depending on the type of infrastructure.

**Keywords:** classification of hyperbolas, Ground penetrating radar, land surveying, the Ring-Projection method, underground infrastructure, urban areas

[1]MSc, Military University of Technology (WAT), Faculty of Civil Engineering and Geodesy, Department of Imagery Intelligence, S. Kaliskiego 2, 00-908 Warsaw, Poland, e-mail: klaudia.pasternak@wat.edu.pl, ORCID: 0000-0001-5794-0560

[2]Associate Prof. PhD, Military University of Technology (WAT), Faculty of Civil Engineering and Geodesy, Department of Imagery Intelligence, S. Kaliskiego 2, 00-908 Warsaw, Poland, e-mail: anna.fryskowska@wat.edu.pl, ORCID: 0000-0001-5449-8092

# 1. Introduction

Research on improving the accuracy and reliability of the non-invasive detection of underground utility (UUs) networks along with their classification based on B-scan images is the key aspect in the field of civil engineering from the point of view of the safety of the conducted infrastructure modernization works that are related to the implementation of BIM technology and 3D cadastre. Non-invasive detection methods make it possible to detect foundations and structures, not visible on building plans, that pose a potential threat to human safety. Accurate information about the location of UUs networks is important in the context of building planning to avoid excavation in the wrong places [1, 2]. Acquiring information on the positions and routes of underground cables and the pipelines has become a necessity in the design and execution processes of all kinds of constructions [3–6].

Automation of the process of detection and extraction of hyperbolas that represent underground objects has been discussed in numerous publications [7–9]. Manual analysis of large amounts of data from ground penetrating radar (GPR) is inefficient and time-consuming. Especially that data received from GPR are images usually named radargrams. Therefore, it is important to develop a fast and accurate method for the automated classification of underground objects. The classification of various types of underground objects in B-scan images is currently limited to mines [10], cavities [11], hatches [12], and soil and underground pipes [13]. The available publications classify underground objects as pipes, without distinguishing between different types of UUs networks. Classification of pipes on echograms can be done by automatic method – using artificial intelligence or by manual (visual) method. In the latter case, it is necessary to analyze the acquired image and determine features representing the underground object, i.e. change in contrast, specific geometric shape. Depending on the type of pipe sought, different features can be used for identification.

In the B-scan images, buried underground objects are represented in the form of hyperbolic patterns. This shape was caused by the reflection of the electromagnetic wave previously sent by the instrument from the object, as recorded by the ground penetrating radar. There are available algorithms that enable the extraction of the selected geometric properties of hyperbolas, that is, the depth, position, and radius of an underground object and its 3D representation [14–17]. However, the use of deep learning in extracting hyperbolas from radargrams has proven to be an effective method for extracting information from radargrams and recognizing hyperbolas using a large dataset of labeled images [18–23]. The results of hyperbola detection and extraction (obtained from deep learning algorithms) are promising, but they have the disadvantage of having to acquire large data sets to be included in the training network. Nowadays, existing algorithms only allow the extraction and detection of underground networks without classifying them according to their type.

Once hyperbolas are extracted from a radargram, they can be used for a variety of purposes. For example, they can be used to detect underground objects, to map the subsurface of an area and to study soil properties. However, images obtained from ground penetrating radars are often highly noise owing to the heterogeneity of the subsoil medium, mutual wave interactions, and hardware module noise, which reduces the efficiency of the extraction

of hyperbolas representing underground objects. The main aim of the filtration of GPR images is to enhance interpretation efficiency and precisely define the position of the objects [24–31]. Radargrams processed in this way are prepared for use in other engineering topics, i.e. tunnel deformation investigation [32] or assessment of bridge condition and location of reinforcement using integrated TLS and GPR data [33].

### 1.1. Proposed method

Summarizing the above related works, our main task was to classify the hyperbolas according to the types of underground utility networks (water supply network "w", sewage network "s", heating network "h", power supply "p", telecommunications "t", and gas networks "g") in images obtained from ground penetrating radars and to automate this process with the use of MATLAB software. In this study, to automatically distinguish utility types effectively, the Ring-Projection method was used to measure signal complexity and irregularity. Then, an approach was proposed that integrates the approach to the classification of types of utility networks based on image and one-dimensional data.

The authors proposed several algorithms to automate the processing of images obtained using the GPR method: algorithm for automated filtration and detection of hyperbolas in radargrams, algorithm for the extraction of hyperbolas and their halves, algorithm for the extraction of geometric properties of hyperbolas and the reduction of 2D data to 1D data and algorithm for automated classification of hyperbolas according to the types of ground utility networks.

## 2. Methodology

In this study, using MATLAB software, combined with the proposed digital image processing algorithm, radargrams were processed by filtering, binarization, and the proposed complex radiometric (based on pixel values and contrast) and geometric (based on the shape and size of objects) features of the hyperbolas algorithm to detect, extract, and classify hyperbolas representing different types of utilities. In general, the proposed method of classifying GPR hyperbolas into different utility types is presented in subsection 2.3.3. The proposed methodology, which is based on the radiometric and geometric information contained in B-scan images. The B-scan transformed to the form of a binary image contains important information that is both radiometric and geometric. The proposed method was based on these properties. However, it should be noted that some radargrams that was also mentioned in this study was impossible to use because of the very poor recording of the image, which made extraction and classification extremely difficult (due to noise, soil layers, roots of trees, and other interfering objects). Therefore, this study focuses primarily on the utility that can be detected.

## 2.1. Data acquisition

The object classification process is preceded by obtaining radargrams, their filtration, and the detection of objects that represent UUs networks. Radargrams were acquired in two research areas: on the campus of the Military University of Technology in Warsaw at 10 different locations and in the city of Otwock at six different locations, using a Leica DS2000 GPR, as presented in Fig. 1.

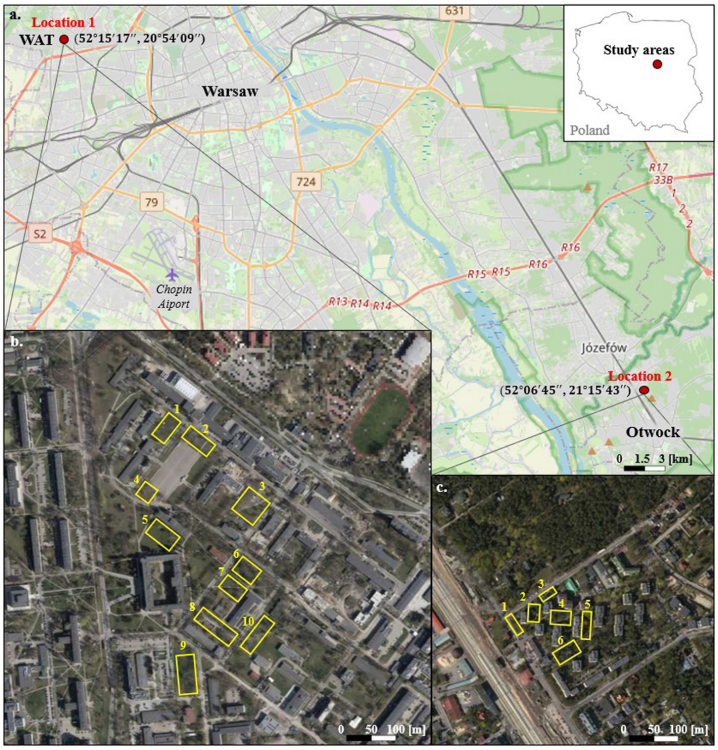


Fig. 1. Location map of the study area showing: (a) the location of measurement areas in the topographic map, (b) Location 1 and (c) Location 2 (Map source: OpenStreetMap, Geoportal)

The measurement areas are marked with yellow rectangles in Fig. 1. In each of these areas, minimum 3 measurement routes were obtained. Radargrams were acquired using GNSS receiver integrated with a ground penetrating radar, which has a positioning accuracy of approximately 1 m. The study used 150 images of true hyperbolas showing all types of UUs networks (about 20–30 images of each type of network). However, it is worth noting that 150 is the number of resultant images of hyperbolas already obtained after the stage of grouping hyperbolas according to their completeness into three groups (Group A, B, and C – described in subchapter 2.3 – Classification). Initially, after the pre-processing of images stage, about 300 images of hyperbolas and other objects detected in the image were obtained.

## 2.2. Pre-processing and processing of images

Pre-processing images is an obligatory stage for acquiring useful GPR radargrams. Radargrams obtained in the uNext software were pre-processed using GRED HD software. This process consists of the stages: time zero correction, bandpass frequency filter, gain filter and background removal.

The algorithm for the detection and extraction of hyperbolas is the first stage of processing and is based on the radiometric properties of the images and objects. This algorithm, which was first proposed in [34], was partly modified for the purposes of hyperbola classification. The new approach is shown in Fig. 2.

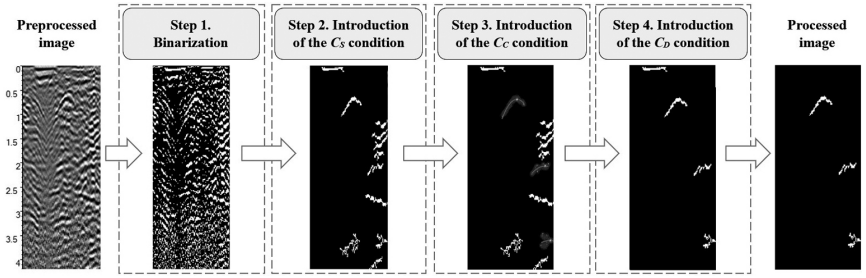


Fig. 2. Scheme of the preliminary research steps (Source: own elaboration)

The stage of edge detection using the Sobel filter [34] was replaced by binarization of the image using the Sauvola method (Step 1 – Binarization). The aim of this action was to preserve not only the edges of the objects but also their overall shape, which is important for the analysis of the geometric properties of the hyperbolas that are subjected to classification. The main steps of the proposed method are as follows:

**Step 1 – Binarization:** The pre-processed image was converted to a gray-scale image by filtering with a mask that removed the background of the image. The second stage consists of the binarization of the image using the Sauvola method [35–37].

**Steps 2-4 – Geometric conditions:** The proposed method for the detection of hyperbolas is based on the extraction of hyperbolas that meet the predefined geometric conditions: ($1^{st}$ condition) the size of the object ($C_S$) is larger than 45 px, ($2^{st}$ condition) the curvature of the object ($C_C$) falls within the range of <0.016; 0.160>, ($3^{st}$ condition) the depth of the object ($C_D$) is larger than 13 px. Entering these conditions means that all hyperbolas not meeting these conditions will be removed from the image. These values were defined based on all the measurement samples (i.e., 20 images).

The resulting binary images obtained as a result of the application of the algorithm (for the detection and extraction of hyperbolas) contained both true and false objects. The term true objects refers to hyperbolas that reliably reflect the position of UUs network according to the National Geodetic and Cartographic Resource (NGCiR). On the other hand, false objects are objects that are detected as true objects but which, in fact, are not ground utility networks. The presence of false objects in the final processed image results from the occurrence of various interfering objects. In comparison to the image in step 1 (Fig. 2), it was possible to eliminate approximately 70–90% of the false objects presented in step 4 (Fig. 2).

## 2.3. Classification

The algorithm for the automated classification of hyperbolas according to the type of ground utility network is mainly based on the determination of the value of their geometric properties. Automated determination of these values requires extracting the hyperbolas and writing them into individual matrices. The resulting grey-scale images contain hyperbolas of various shapes and degrees of completeness. The authors proposed dividing hyperbolas into three groups according to the degree of completeness:

– complete objects (with a vertex and both arms of a hyperbola) – Group A,
– objects that are approximately complete (with the vertex and left or right arm of the hyperbola) – Group B,
– incomplete objects (with only one arm or vertex of the hyperbola) – Group C.

Only hyperbolas from groups A and B were selected for the analysis of the geometric properties. Hyperbolas from group C should be rejected and should not be considered in the classification process because this group consists of incomplete objects that do not have a hyperbolic shape. Fig. 3 shows the true hyperbolas representing all types of UUs networks analysed.

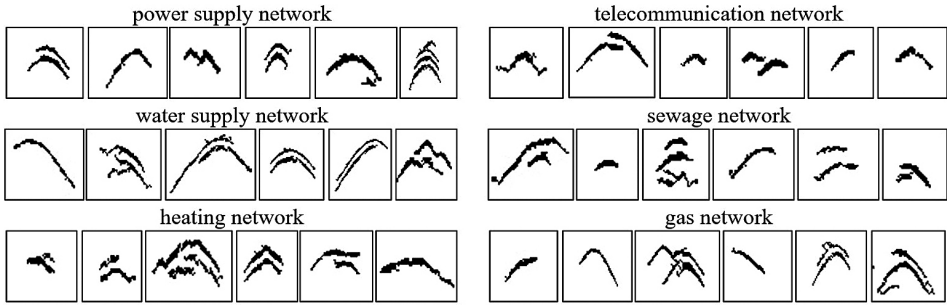


Fig. 3. Grey-scale images of hyperbolas representing various types of utility networks (Source: own elaboration)

The analysis in Fig. 3 reveals that the visual classification of ground utility network types requires analytical experience. Therefore, in the following steps, the geometric properties of these types are presented, as well as the possibility of their dimensioning for semi-automated classification of utility network types.

### 2.3.1. Extraction of hyperbolas

In order to reliably compare the geometric properties of the hyperbolas in Group A and B, the images of the halves of the hyperbolas were analyzed (due to the symmetry of the hyperbolas and their degree of completeness). For hyperbolas the coordinates of the borderline points of each object were determined (marked with red circles in Fig. 4) and defined the vertex (marked with a green circle in Fig. 4). This will be the basis for automated cutting out of two images containing halves of the hyperbola.

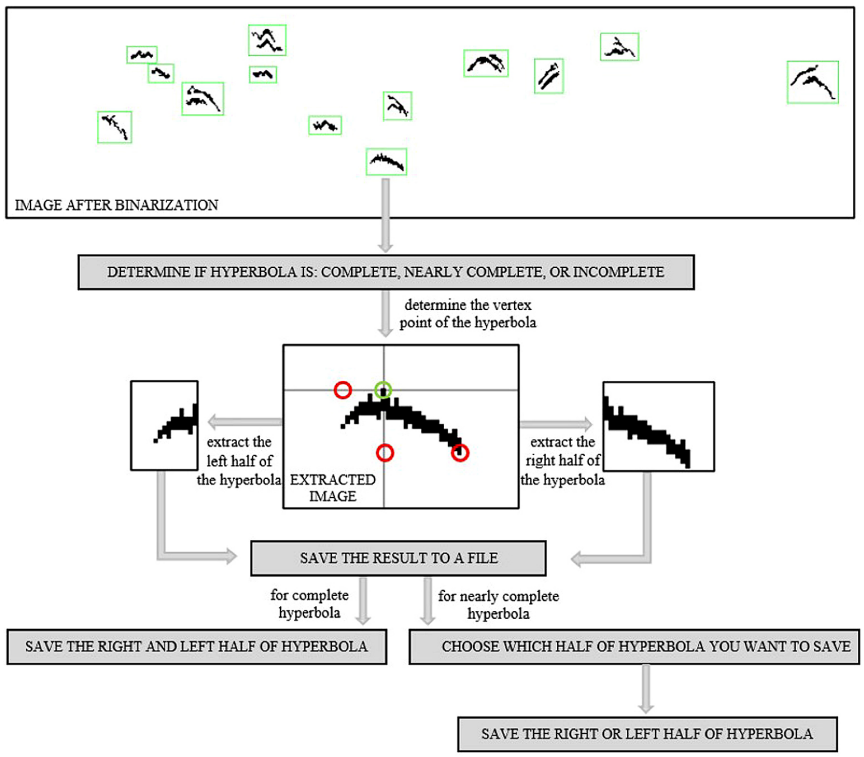


Fig. 4. Flowchart of the algorithm for the automated extraction of hyperbolas and cutting their halves out of the resultant binary image (Source: own elaboration)

For hyperbolas belonging to Group A, two images representing both halves were saved. In group B, the user selects which half of the hyperbola should be saved.

### 2.3.2. Geometric features of hyperbolas. proposed parameters and coefficient

The geometric features of the hyperbolas represent their geometric properties in the image. The process of measuring the hyperbola enables to obtain the numerical characteristics of the analyzed object. In further steps, we propose to determine the value of the following properties of the hyperbolas: curvature ($C_C$), span ($R$) and height ($H$), size ($C_S$), depth ($C_D$), signal length ($L_S$), signal height ($H_S$). The determined values of the geometric properties of the hyperbolas were the basis for calculating the f parameter, which is the ratio of the ratio of the product of the power of the span ($R$) to sixteen times the height of the hyperbola ($H$), the $K$ coefficient, which is the ratio of the height ($H$) to the span ($R$) of the hyperbola, and $ODL$ is the distance between the vertex and the end of the arm of the hyperbola. These parameters were calculated using Eqs. (2.1), (2.2) and (2.3).

$$f = \frac{2R^2}{16H} \tag{2.1}$$

$$K = \frac{H}{R} \tag{2.2}$$

$$ODL = \sqrt{H^2 + R^2} \tag{2.3}$$

where: $H$ – the hyperbola height ($H_L$ – for the left half of the hyperbola, $H_R$ – for the right half of the hyperbola) [px], $R$ – the hyperbola span ($R_L$ – for the left half of the hyperbola, $R_R$ – for the right half of the hyperbola) [px].

For hyperbolas belonging to group A, the values of height and span were calculated for both the left ($R_L$, $H_L$) and right ($R_R$, $H_R$) halves. The values were averaged to calculate the $K$. For hyperbolas from group B, these parameters were calculated for only one-half. All the properties discussed so far were determined based on the image of the whole hyperbola or one of its halves. However, the length ($L_S$) and height ($H_S$) of the signal were calculated based on one-dimensional data that represented half of the hyperbolas from groups A and B. These data were obtained by reducing the image (2D) of the hyperbola to the signal (1D) using the Ring-Projection-Wavelet-Fractal method, which is discussed in subchapter 2.3.2.1.

**2.3.2.1. The Ring-Projection method**

In this method, the first step to reduce the dimension from 2D to 1D is to determine the center of mass ($M_O$) for each object using a density function of the mass distribution on the plane. On this basis, Cartesian coordinates were converted into a polar system (from $x$–$y$ space to $\gamma - \theta$ space) (5) [38]. Then, rings with radii ($r + \Delta r$) centered at point $M_O$ were created. In the last step of the algorithm, the number of pixels in the range of the given ring is counted for each value of radius $r$. Fig. 5 shows a diagram of the reduction of the image dimension to one-dimensional data on the example of a selected hyperbola, which is a power supply network.

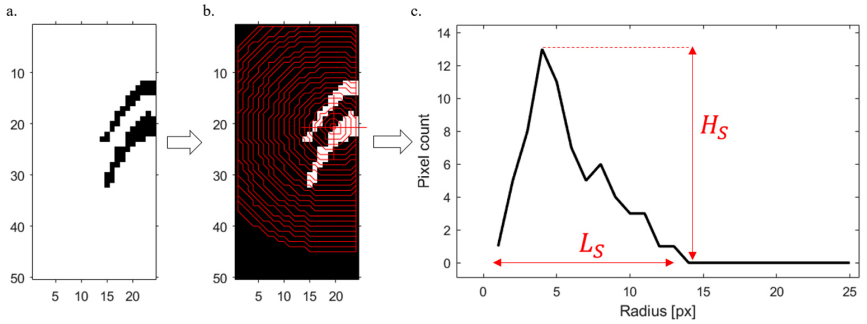


Fig. 5. Scheme of the reduction of image dimension to signal: (a) binary image of a hyperbola, (b) centroid of the mass distribution on the $x$–$y$ plane with rings of a given radius, and (c) ring projection of hyperbola image (Source: own elaboration based on [37, 38])

This method was applied to avoid considering the disturbances in the extraction of the $L_S$ and $H_S$ parameters of true hyperbolas considering the potential reduction or expansion of the shapes of complete and nearly complete hyperbolas. In this way, these parameters are resistant to varying hyperbola lengths or curvatures.

### 2.3.3. Classification method

The processed images, extracted hyperbolas and their halves were the basis for developing a method of classifying objects according to the types of ground utility networks based on the extraction of their geometric properties. The values of the $C_C$, $C_D$ parameters were obtained from images of hyperbolas, whereas the values of the parameters: $R$, $H$, $C_S$, $H_S$, $L_S$ were obtained from images of their halves. The proposed method allows for the classification of hyperbolas based on the calculated values of the parameters and coefficients ($f$, $K$, $ODL$). The stages of manual work are marked with red dotted lines (Fig. 6).

The correctness of the classification of utility networks was verified based on reference data and is described in the Results section. Twelve main conditions for the classification of hyperbolas were introduced, and their ranges were determined heuristically based on the test hyperbolas.

1. The power supply network was classified as "p", if: the value of the $C_D$ parameter falls into the range $\ll 24; 44 \gg$ [px].
2. The water supply network was classified as "w" if: the value of the $C_S$ parameter falls into the range $\ll 120; 190 \gg$ [px] or the value of the $ODL$ parameter <53 [px] or the value of the $C_S \geq 60$ px and of $L_S \geq 19$ px or the value of the $f$ parameter falls into the range $\ll 5.1; 6.0 \gg \wedge \ll 6.4; 8.4 \gg$.
3. The sewage network was classified as "s" if: the value of the $C_C$ parameter falls into the range $\ll 0.020; 0.021 \gg$ [px] or the $ODL$ parameter falls into the range <36; 51>[px] or the value of the $C_S$ parameter falls into the range $\ll 115; 200 \gg$ [px].
4. The telecommunications network was classified as "t" if: the value of the $C_S$ parameter falls into the range $< 24; 35 >$ [px] or the $ODL$ parameter falls into the range $\ll 15; 17 \gg$ [px] and the value of $C_D \leq 19$ px.
5. The heating network was classified as "h" if: the value of $C_S > 60$ px and $L_S < 19$ px.

These conditions allow the classification of the following types of networks: p, t, s, h and w. As a result of the classification of hyperbolas, the power supply and heating networks may be classified (i.e. meet the adequately defined number of conditions): once, the water supply network four times at the maximum, and the sewage network three times. The results of the preliminary classification can be analyzed for several variants (cases):

**Case 1.** The result of the classification is a specific type of network, for example w, s, h, t, or p. (one condition has been met).

**Case 2.** The result of the classification is that more than one type of network, which means that more than one condition has been met. If at least two results indicate one type of network, then the hyperbola should be classified as the dominant type of network, for example: "ssw" (3 conditions met) → "s", "ssww" (4 conditions met) → "sw", "ppt" (3 conditions met) → "p" and "hhww" (4 conditions met) → "hw".

Thus, the final result in Case 2 may be a single (e.g. "w", "s", "h", "t" or "p") or double network (e.g. "sw", "hw" or "pt").

**Case 2.1.** If, in Case 2, the result is "pt" then additional verification should be conducted: The power supply network was classified as "p", if:

– (single hyperbola), the value of the $C_D$ parameter fell into the range $< 24; 34 >$ and the value of the $C_S$ parameter fell into the range $< 24; 30 >$,

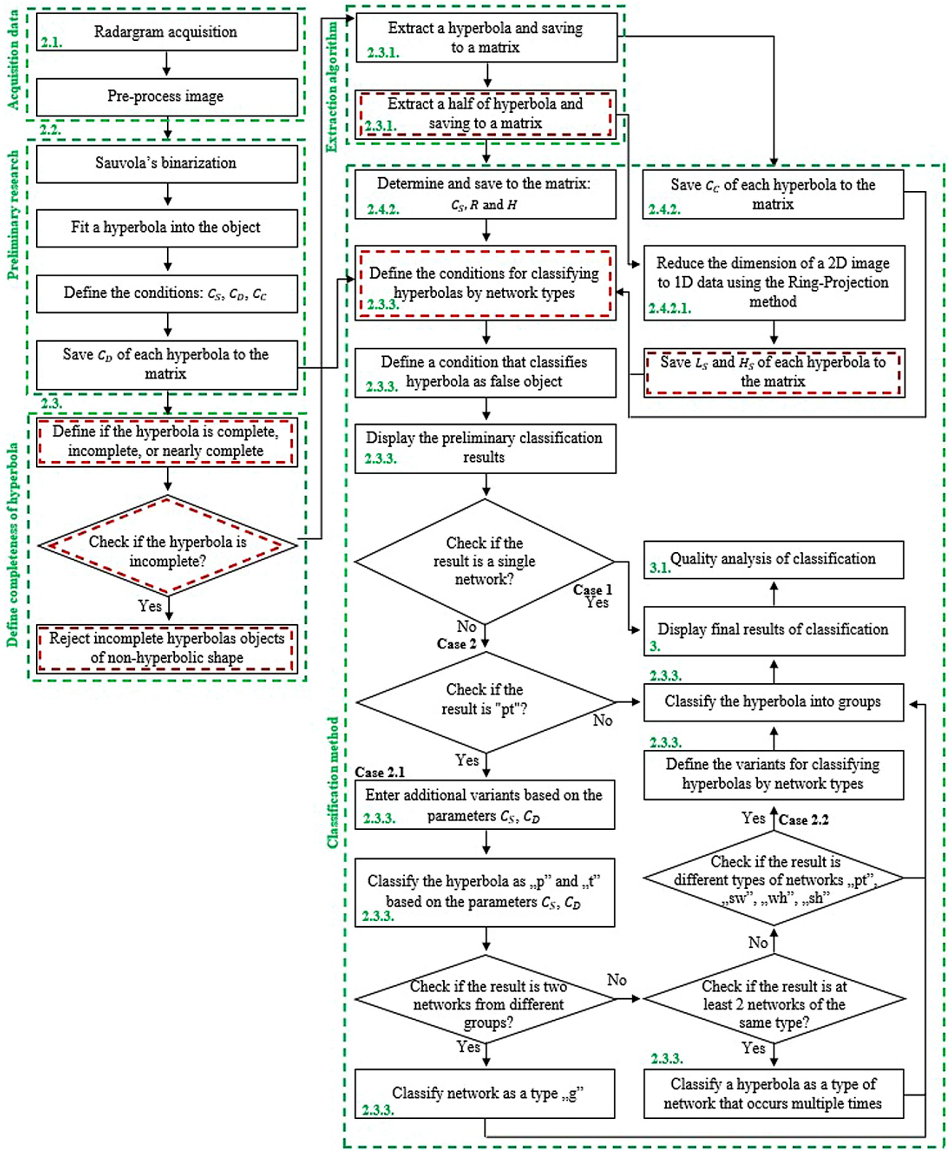


Fig. 6. Schematic diagram of the classification algorithm of hyperbolas (Source: own elaboration)

– (double hyperbola), the value of the $C_S$ parameter fell into the range $< 24; 34 >$ and the value of the $C_D$ parameter fell into the range $< 24; 34 >$ for one object and in the range $< 34; 44 >$ for the second object.

The telecommunications network was classified as "t" if:

– (single hyperbola), the value of the $C_D$ parameter fell into the range $< 34; 44 >$ and the value of the $C_S$ parameter fell into the range $< 24; 30 >$.

The additional verification presented above allowed us to partly separate the single "p" or "t" networks from the double "pt" results.

**Case 2.2.** If, in Case 2.1, the obtained result is still "pt" or other instances of double network ("sw", "hw" or "sh"), the additional conditions presented below should be applied:

- The power supply network should be classified as "p" if $f < 5.0 \wedge C_{S1} > 56$ px and $H_S \geq 10$,
- The telecommunications network should be classified as "t" if $f < 5.0 \wedge C_{S1} \leq 56$ px and $H_S < 10$,
- The heating network should be classified as "h" if $f \leq 3.6 \wedge C_{S1} > 50$ px and $H/R > 0.8$
- The water supply network should be classified as "w" if $f > 5.0 \wedge C_{S1} > 50$ px and $H/R > 0.8$,
- The sewage network should be classified as "s" if $f < 5.0 \wedge C_{S1} < 50$ px and $H/R < 0.8$.

The final result of the classification should be one or, at maximum, two networks. The next stage of classification consisted of assigning the given networks to three groups. The given hyperbola is classified into one of three groups based on the physical properties of the analyzed objects, that is, diameter, thickness, depth at which the network is situated, and curvature of the hyperbola. The first group (Group 1) includes networks that are in the form of underground cables and, at the same time, are located at a more shallow depth than those from the second group (Group 2), which are underground pipes and contain liquid materials inside. These networks were characterized by a similar range of diameters. Finally, the third group (Group 3) includes only one network with the properties of both the first and second groups. The networks were assigned to the defined groups as follows:

- Group 1: power supply ("p") and telecommunication ("t") networks,
- Group 2: water supply ("w"), heating ("h") and sewage ("s") networks,
- Group 3: gas network ("g").

Based on an analysis of 150 images of hyperbolas, it was found that the gas network is the only utility network that does not have the characteristic properties specified above. It was classified, if the result of the classification were two networks that belonged to different groups (1–3), e.g. "pw" or "ph" due to the presence of the characteristic features of networks from groups 1 and 2.

## 3. Results and quality analysis

This section presents the results of the classification of hyperbolas according to the types of ground utility networks along with an assessment of the quality of this classification. The obtained classification results were verified based on reference data obtained from the National Geodetic and Cartographic Resource (NGCiR). Fig. 7 presents the classification results considering the types of networks in the given class (Fig. 7a) and the networks classified as the given type in other classes (Fig. 7b).

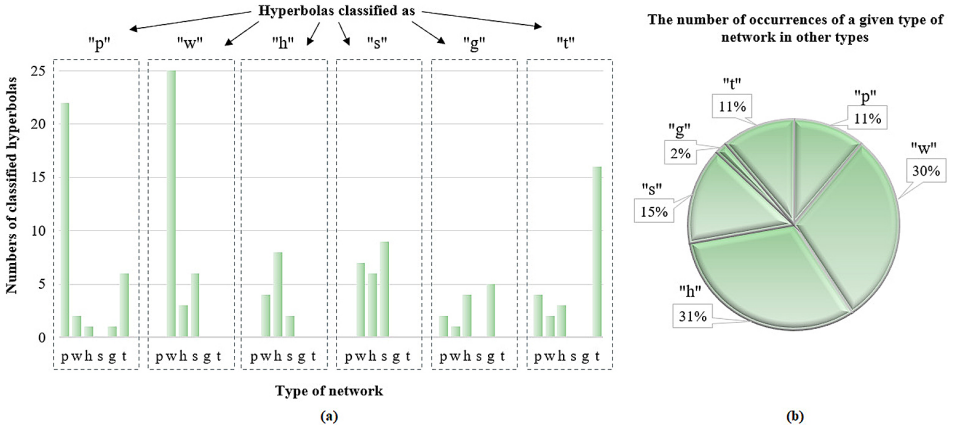


Fig. 7. The results of the classification of hyperbolas (a) by network types in a given class, (b) classified as a given type in other classes (Source: own elaboration)

The results of the classification of the hyperbola data enabled the assignment of all classified hyperbolas to network groups (I–III). The effectiveness of the assignment to the given group – $R_g$, was calculated, and the results of the average effectiveness of the classification are presented in Table 1. The overall average effectiveness of the classification of hyperbolas according to the network groups (I–III) was 77%.

Table 1. The effectiveness of the classification of hyperbolas according to network groups (I–III)

| Group of networks | I | II | III |
|---|---|---|---|
| $\overline{R_g}$ [%] | 82 | 86 | 62 |

The result of the classification may be one (case 1) or two networks (case 2) belonging to the same group of networks (I–III). The user who employs the proposed algorithm for the automated classification of network types will receive the final result of classification of the given hyperbola in the form of the effectiveness of classification $R_C1$ expressed in the form of a percentage (Eq. 3.1). This effectiveness considers the effectiveness of classification of the given type of network $R_s$ and the average effectiveness of assigning the hyperbola to the given group (I–III).

$$R_{C1} = \frac{R_s + R_g}{2} \tag{3.1}$$

where: $R_s$ – the effectiveness of classification of the given type of network, $R_g$ – the effectiveness of the assignment to the given group (I–III), $R_C1$ – final effectiveness of classification considering the effectiveness of: the assignment to the given group ($R_g$) and of classification of the given type of network ($R_s$).

The classification result $R_s$ may take the values 0, 33, 50%, or 100%, depending on the number and type of networks assigned to the given hyperbola. In the first case (if the classification result is: “p”, “t”, “g”, “h”, “w” or “s”), the effectiveness of classification

of the given type of network may be either 0 or 100%. In the second case, however (if the classification result is: "pt", "sh", "sw" or "wh") all networks are characterized by a 50% classification effectiveness. Table 2 presents the average results of the effectiveness of the classification of hyperbolas according to ground utility network types – $\overline{R_s}$ and the effectiveness of this classification, considering the effectiveness of the classification of the given hyperbolas to groups (I–III) – $R_{sg}$.

Table 2. Average results of the efficiency of the classification of hyperbolas according to types of ground utility networks

| Classification type | Result of the classification of hyperbolas according to types of utility networks | | | | | |
|---|---|---|---|---|---|---|
| | p | w | h | s | g | t |
| $\overline{R_s}$[%] | 60 | 56 | 30 | 38 | 46 | 55 |
| $\overline{R_{sg}}$ [%] | 63 | 70 | 53 | 49 | 42 | 54 |

As a result of the application of the proposed object classification method, not all hyperbolas representing ground utility networks were classified. Some of these did not meet any of the proposed conditions. Table 3 presents the percentage of hyperbolas that were not classified according to the types of networks – ($N_s$).

Table 3. Results for unclassified hyperbolas by type of network – $N_s$

| Type of network | p | w | h | s | g | t |
|---|---|---|---|---|---|---|
| $N_s$ [%] | 3 | 24 | 33 | 29 | 25 | 14 |

Table 4 presents the percentage of hyperbolas that were not classified according to their group (I–III) – $N_{sg}$.

Table 4. Percentage of hyperbolas that were not classified according to groups (I–III) – $N_{sg}$

| Group of networks | I | II | III |
|---|---|---|---|
| $N_{sg}$ [%] | 9 | 29 | 25 |

The binary images obtained as a result of the image-processing stage contain images of both true and false hyperbolas. The effectiveness of detecting these hyperbolas was tested on 10 binary images. True objects accounted, on average, for approximately 60%, and false objects for approximately 40% of all objects detected in the images. The false objects identified in the images also participated in the process of classifying the ground utility network types. On an average, half of these false objects were classified as a type of network. The process of classifying false objects was tested using ten images. Among them, the minimum number of classified false objects was 27% and the maximum was 75%.

## 3.1. Classification quality

Accuracy, Recall, Precision and F-measure were used to evaluate the classification performance. To calculate these metrics, the values of the True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN) parameters should be defined for each type of network based on the reference data and the results obtained in the classification of hyperbolas. Based on the defined confusion matrices for each type of network, the results of the classification of hyperbolas were eva luated by determining the values of the following parameters: Recall, Precision, F-measure, and Accuracy [39, 40] for each type of network (Table 5).

Table 5. Calculated metrics: Recall, Precision, F-measure and Accuracy, for all types of UUs networks

| Type of network | p | w | h | s | g | t |
|---|---|---|---|---|---|---|
| Recall | 0.7097 | 0.7429 | 0.5714 | 0.4286 | 0.4167 | 0.6400 |
| Precision | 0.7586 | 0.6341 | 0.3333 | 0.5294 | 0.8333 | 0.7273 |
| F – measure | 0.7333 | 0.6842 | 0.4211 | 0.4737 | 0.5556 | 0.6809 |
| Accuracy | 0.8416 | 0.7818 | 0.7944 | 0.8095 | 0.9140 | 0.8500 |

If a high value of the Recall parameter and a low value of Precision are obtained, one may conclude that the model classifies most positive hyperbolas correctly, but, at the same time, brings many false positive results, i.e. classifies many negative hyperbolas as positive ones. The described case refers to the Recall and Precision parameters obtaine d for hyperbolas belonging to the heating network. The highest classification (Accuracy) was achieved for the classification of hyperbolas belonging to the gas network (0.9140), telecommunications network (0.8500), and power supply network (0.8416). The highest values of the calculated metrics (Recall, Precision, F-measure and Accuracy) for the analyzed types of networks were achieved for power supply, telecommunications, and water supply networks.

## 3.2. Effectiveness and Quality factor ($F_{EQ}$)

Considering the results of the assessment of the accuracy of the classification of types of hyperbolas, the authors propose to int roduce a coefficient to evaluate the classification (the Effectiveness and Quality factor), which will determine the effectiveness by analyzing various instances of results for each type of network (Eqs. 3.2–3.7). For example, the power supply network in our results occurred in cases: “p” and “pt”. Due to that, $F_{EQ}$ for the power supply network was calculated based on the weighted average, taking into account the average effectiveness of classification of the power supply network for cases “p” and “pt” and the number of occurring cases for “p” and “pt”. This factor considers only the classification results obtained for true hyperbolas.

$$F_{EQ_{p,pt}} = \frac{n_p \cdot \overline{R_{sg_p}} + n_{pt} \cdot \overline{R_{sg_{pt}}}}{n_p + n_{pt}} \cong 78\% \tag{3.2}$$

$$F_{EQ_{t,pt}} = \frac{n_t \cdot \overline{R_{sg_t}} + n_{pt} \cdot \overline{R_{sg_{pt}}}}{n_t + n_{pt}} \cong 78\% \tag{3.3}$$

$$F_{EQ_{w,wh,sw}} = \frac{n_w \cdot \overline{R_{sg_w}} + n_{wh} \cdot \overline{R_{sg_{wh}}} + n_{sw} \cdot \overline{R_{sg_{sw}}}}{n_w + n_{wh} + n_{sw}} \cong 72\% \tag{3.4}$$

$$F_{EQ_{h,wh,sh}} = \frac{n_h \cdot \overline{R_{sg_w}} + n_{wh} \cdot \overline{R_{sg_{wh}}} + n_{sh} \cdot \overline{R_{sg_{sh}}}}{n_h + n_{wh} + n_{sh}} \cong 59\% \tag{3.5}$$

$$F_{EQ_{s,sw,sh}} = \frac{n_s \cdot \overline{R_{sg_s}} + n_{sw} \cdot \overline{R_{sg_{sw}}} + n_{sh} \cdot \overline{R_{sg_{sh}}}}{n_s + n_{sw} + n_{sh}} \cong 72\% \tag{3.6}$$

$$F_{EQ_g} = \frac{n_g \cdot \overline{R_{sg_g}}}{n_g} \cong 81\% \tag{3.7}$$

where: $n_i$ – is the number of occurrences of the given case, where $i$ = p, pt, t, w, wh, sw, h, sh, g, $R_{sgi}$ – is the average effectiveness of the network classification for the occurrence of the given instance $i$.

The proposed methodology was implemented in MATLAB software, which enables the display of individual stages of the work, and thus the qualitative analysis of the classification. As a result of the application of the method, the analyst receives information about the extracted hyperbolas, their classification, and the evaluation of the quality of the classification in the form of: types of the classified ground utility networks, average effectiveness of classification of the given type of network that considers the effectiveness of assigning the hyperbola to the given group (I–III) – $R_{sg}$ and $F_{EQ}$. Fig. 8 presents the interface of the developed algorithm, as seen by the user. The data that will be displayed to the user as a result of the operation of the algorithm are $R_C1$ and $F_{EQ}$.

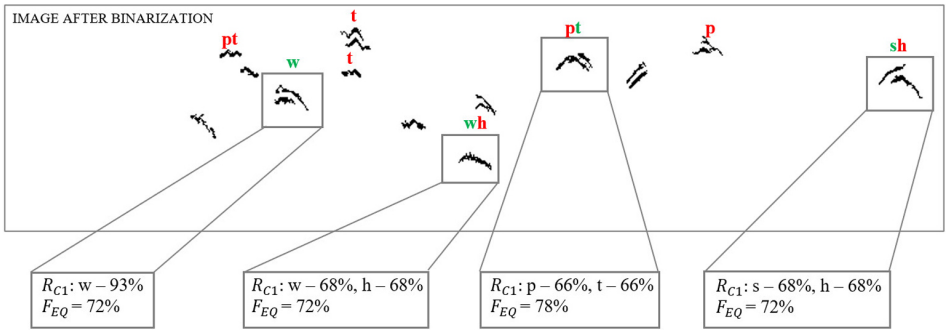


Fig. 8. Interface of the algorithm for the classification of network types as seen by the user (Source: own elaboration)

The proposed algorithm for automated classification of hyperbolas was tested on 12 complete measurement routes, from which approximately 150 images of hyperbolas representing ground utility networks were extracted.

# 4. Conclusions, limitations and future works

The subject literature review revealed several problems concerning the automated processing of GPR data for extracting specific information. The two major knowledge gaps are the automated classification of UUs network types and a clear assessment of the quality of such classifications. This motivated us to develop the proposed approach. The algorithm for classifying UUs networks operates in a semi-automatic mode. As a result of the operation of the algorithm, the user receives information about the type of classified UUs network, the average effectiveness of classification of the type of network considering the classification of network groups, the overall accuracy of classification, the $F_{EQ}$ factor, as well as the commission and omission errors. These data improve the classification process and support operators, especially inexperienced ones, in making decisions regarding the type of ground utility network.

The developed methodology allows the classification of each network type. The highest classification effectiveness $\overline{R_s}$ (60%) was noted for the power supply network, whereas the lowest effectiveness (30%) was achieved for the heating network. However, considering the assignment of the given hyperbola to a group of networks (I–III), the effectiveness of the classification $\overline{R_{sg}}$ increased significantly, ranging from 42% (for the classification of the gas network) to 70% (for the water supply network). Based on the classification results, the authors proposed introducing a coefficient of the quality of classification ($F_{EQ}$). This factor was introduced because of the occurrence of different instances of network classification in the final results. The highest value of the factor was achieved for the classification of the gas network (81%), whereas the lowest value (59%) was noted for the heating network. This value ($F_{EQh,wh,sh}$) was influenced by the co-existence of the target network type with another type of network.

Compared to conventional and existing methods, the degree of automation of the proposed method is significantly higher, which enables the use of the original images obtained from GPR as input data with very little manual intervention. Moreover, the proposed algorithm for the extraction of hyperbolas of various shapes and degrees of completeness makes it possible to analyze a significantly higher number of hyperbolas, and is resistant to the quality of images, which varies depending on the conditions of obtaining data. The entire process is highly automated, and as such, it saves time and cost associated with manual analysis and possible corrections. Although the effectiveness of the proposed methodology has been verified in this study, it still suffers from some limitations, including: manual vertex extraction and manual verification of completeness. These limitations may introduce some uncertainty to the classification. According to the authors, these steps can now be more precisely performed manually by any even inexperienced operator.

The proposed novel method for classifying types of UUs networks from radargrams greatly automates and significantly speeds up the process of network classification, which is a major contribution to research in this topic. This is evidenced by the proposed parameters and conditions (2.2–2.3) based on geometric features, which allowed to characterize and distinguish the given network types. Further research on this issue will concern both the issue mentioned above and the limitations of improving the effectiveness of the extraction of true hyperbolas.

## References


[1] P. Stones and T.Y. Heng, "Underground space development key planning factors", *Procedia Engineering*, vol. 165, pp. 343–354, 2016, doi: 10.1016/j.proeng.2016.11.709.

[2] C. Akcay, A. Solt, and N.M. Korkmaz, "A proposal for the reconstruction of a historical masonry building constructed in Ottoman Era (Istanbul)", *Journal of Building Engineering*, vol. 32, art. no. 101493, 2020, doi: 10.1016/j.jobe.2020.101493.

[3] M. Metwaly, "Application of GPR technique for subsurface utility mapping: A case study from urban area of Holy Mecca, Saudi Arabia", *Measurement*, vol. 60, pp. 139–145, 2015, doi: 10.1016/j.measurement.2014.09.064.

[4] W. Nawrocki and Z. Piasek, "Experimental testing of-geo-radar resolving power for detection of underground installations", *Geodesy and Cartography*, vol. 54, no. 3, pp. 119–130, 2005.

[5] S. Cafiso, A. Di Graziano, D. Goulias, M. Mangiameli, and G. Mussumeci, "Implementation of GPR and TLS data for the assessment of thepp. bridge slab geometry and reinforcement", *Archives of Civil Engineering*, vol. 66, no. 1, pp. 297–308, 2020, doi: 10.24425/ace.2020.131789.

[6] A. Hoła, Ł. Sadowski, and J. Szymanowski, "Non-destructive testing and analysis of a XIX Century Brick Masonry Building", *Archives of Civil Engineering*, vol. 66, no. 4, pp. 201–219, 2020, doi: 10.24425/ace.2020.135217.

[7] S. Birkenfeld, "Automatic detection of reflexion hyperbolas in GPR data with neural networks", in *World Automation Congress*. Kobe, Japan, 2010, pp. 1–6.

[8] H. Liu, C. Lin, J. Cui, L. Fan, X. Xie, and B. F. Spencer, "Detection and localization of rebar in concrete by deep learning using ground penetrating radar", *Automation in Construction*, vol. 118, art. no. 103279, 2020, doi: 10.1016/j.autcon.2020.103279.

[9] W. Lei, F. Hou, J. Xi, Q. Tan, M. Xu, X. Jiang, G. Liu, and Q. Gu, "Automatic hyperbola detection and fitting in GPR B-scan image", *Automation in Construction*, vol. 106, art. no. 102839, 2019, doi: 10.1016/j.autcon.2019.102839.

[10] U.S. Khan, W. Al-Nuaimy, and F.E. Abd El-Samie, "Detection of landmines and underground utilities from acoustic and GPR images with a cepstral approach", *Journal of Visual Communication and Image Representation*, vol. 21, no. 7, pp. 731–740, 2010, doi: 10.1016/j.jvcir.2010.05.007.

[11] M.S. Kang, N. Kim, S.B. Im, J.-J. Lee, and Y.-K. An, "3D GPR image-based UcNet for enhancing underground cavity detectability", *Remote Sensing*, vol. 11, no. 21, 2019, doi: 10.3390/rs11212545.

[12] N. Kim, S. Kim, Y.K. An, and J.J. Lee, "Triplanar imagsubing of 3-D GPR data for deep-learning-based underground object detection", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 11, pp. 4446–4456, 2019, doi: 10.1109/JSTARS.2019.2953505.

[13] N. Kim, S. Kim, Y.K. An, and J.J. Lee, "A novel 3D GPR image arrangement for deep learning-based underground object classification", *International Journal of Pavement Engineering*, vol. 22, no. 6, pp. 740–751, 2021, doi: 10.1080/10298436.2019.1645846.

[14] R. Parrillo, R.L. Roberts, and A. Haggan, "Bridge deck condition assessment using ground penetrating radar", presented at International Bridge Conference, Pittsburgh, PA, 2005.

[15] L.L. olde Scholtenhuis, X. den Duijn, and S. Zlatanova, "Representing geographical uncertainties of utility location data in 3D", *Automation in Construction*, vol. 96, pp. 483–493, 2018, doi: 10.1016/j.autcon.2018.09.012.

[16] B. Jafrasteh and N. Fathianpour, "Automatic extraction of geometrical characteristics hidden in ground-penetrating radar sectional images using simultaneous perturbation artificial bee colony algorithm", *Geophysical Prospecting*, vol. 65, no. 1, pp. 324–336, 2017, doi: 10.1111/1365-2478.12413.

[17] J. Gocał, Ł. Ortyl, and M. Sołtys, "Processing of images recorded by the radar method, aimed at construction of metric 3D models of subsurface objects and structures", *Geodesy and Cartography*, vol. 55, no. 1, pp. 47–56, 2006.

[18] U. Ozkaya, F. Melgani, M. Belete, B. Levent, and S.M. Donelli, "GPR Bscan image analysis with deep learning methods", *Measurement*, vol. 165, art. no. 107770, 2020, doi: 10.1016/j.measurement.2020.107770.

[19] P. Gamba, "Neural detection of pipe signatures in GPR images", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 790–797, 2000, doi: 10.1109/36.842008.

[20] Z.J. Long, B.A. Xing, H. Liu, et al., "Hyperbola recognition from ground penetrating radar using deep convolutional neural networks", in *2nd International Conference on Artificial Intelligence: Techniques and Applications*. Shenzhen, China, 2017, pp. 11–13.

[21] M. Wlodarczyk-Sielicka, J. Lubczonek, and A. Stateczny, "Comparison of selected clustering algorithms of raw data obtained by interferometric methods using artificial neural networks", in *Proceedings of the 17th International Radar Symposium.* Krakow, Poland, 2016, pp. 1–5.

[22] A. Czechowicz, P. Gryboś, J. Jachimski, S. Mikrut, Z. Mikrut, P. Pawlik, and R. Tadeusiewicz, *Sieci neuronowe w procesach dopasowania zdjęć lotniczych, (Neural networks in aerial images matching process)*. AGH Krakow, 2010 (in Polish).

[23] A. Sekrecka, M. Kedzierski, and D. Wierzbicki, "Pre-Processing of panchromatic images to improve object detection in pansharpened images", *Sensors*, vol. 19, no. 23, 2019,sub doi: 10.3390/s19235146.

[24] J. Lester and L.E. Bernold, "Innovative process to characterize buried utilities using Ground Penetrating Radar", *Automation in Construction*, vol. 16, no. 4, pp. 546–555, 2007, doi: 10.1016/j.autcon.2006.09.004.

[25] X. He, C. Wang, R. Zheng, and X. Li, "GPR image noise removal using Grey Wolf Optimisation in the NSST Domain", *Remote Sensing*, vol. 13, no. 21, 2021, doi: 10.3390/rs13214416.

[26] Z. Xiang, G. Ou, and A. Rashidi, "Robust cascaded frequency filters to recognize rebar in GPR data with complex signal interference", *Automation in Construction*, vol. 124, art. no. 103593, 2021, doi: 10.1016/j.autcon.2021.103593.

[27] X. He, H. Yan, C. Wang, R. Zheng, Y. Li, and X. Li, "Non-stationary random noise removal in ground-penetrating radar images subby using self-guided filtering", *Digital Signal Processing*, vol. 129, art. no. 103690, 2022, doi: 10.1016/j.dsp.2022.103690.

[28] L. Arévalo-Lomas, B. Biosca, D. Paredes-Palacios, and J. Díaz-Curiel, "Processing radargrams to obtain resistivity sections", *Remote Sensing*, vol. 14, no. 11, 2022, doi: 10.3390/rs14112639.

[29] L. Chen and Y. Wang, "Automatic key frame extraction in continuous videos from construction monitoring by using color, texture, and gradient features", *Automation in Construction*, vol. 81, pp. 355–368, 2017, doi: 10.1016/j.autcon.2017.04.004.

[30] Z. Hui-lin, T. Mao, and C. Xiao-li, "Feature extraction and classification of echo signal of ground penetrating radar", *Wuhan University Journal of Natural Sciences*, vol. 10, pp. 1009–1012, 2005, doi: 10.1007/BF02832458.

[31] U. Marmol and N. Borowiec, "Detection of line objects by means of Gabor Wavelets and Hough transform", *Archives of Civil Engineering*, vol. 66, no. 3, pp. 339–363, 2020, doi: 10.24425/ace.2020.134401.

[32] A. Lejzerowicz and M. Wutke, "GPR as a support tool in determining the causes of failure of objects built in the "White Box" technology", *Archives of Civil Engineering*, vol. 66, no. 3, pp. 173–189, 2020, doi: 10.24425/ace.2020.134391.

[33] S. Cafiso, A. Di Graziano, et al., "Implementation of GPR and TLS data for the assessment of the bridge slab geometry and reinforcement", *Archives of Civil Engineering*, vol. 66, no. 1, pp. 297–308, 2020, doi: 10.24425/ace.2020.131789.

[34] K. Onyszko and A. Fryśkowska-Skibniewska, "A new methodology for the detection and extraction of hyperbolas in GPR images", *Remote Sensing*, vol. 13, no. 23, 2021, doi: 10.3390/rs13234892.

[35] A. Zhang, K.C.P. Wang, E. Yang, J.Q. Li, C. Chen, and Y. Qiu, "Pavement lane marking detection using matched filter", *Measurement*, vol. 130, pp. 105–117, 2018, doi: 10.1016/j.measurement.2018.07.089.

[36] L.D. Sharma and R.K. Sunkaria, "A robust QRS detection using novel pre-processing techniques and kurtosis based enhanced efficiency", *Measurement*, vol. 87, pp. 194–204, 2016, doi: 10.1016/j.measurement.2016.03.015.

[37] K. Onyszko and A. Fryśkowska-Skibniewska, "Analysis of the possibility of classification the types of utilities networks on the GPR images", presented at FIG Congress - Volunteering for the future - Geospatial excellence for a better living, 11–15 September 2022, Warsaw, Poland, 2022.

[38] P. Guan-Chen, "A tutorial of wavelet for pattern recognition", National Taiwan University, Taipei, Taiwan, 2011.

[39] M. Łącka and J. Łubczonek, "Analysis of qualitative and quantitative assessment methods for shoreline extraction", *Zeszyty Naukowe Akademii Morskiej w Szczecinie*, no. 69, pp. 9–19, 2022, doi: 10.17402/495.

[40] W. Lu, J. Dong, et al., "Damage identification of bridge structure model based on empirical mode decomposition algorithm and Autoregressive Integrated Moving Average procedure", *Archives of Civil Engineering*, vol. 68, no. 4, pp. 653–667, 2022, doi: 10.24425/ace.2022.143060.

# Automatyczna klasyfikacja podziemnych sieci uzbrojenia terenu na obszarach zurbanizowanych: nowatorska metoda klasyfikacji integrująca detekcję GPR i przetwarzanie obrazów

**Słowa kluczowe:** georadar, klasyfikacja typów sieci, podziemne sieci uzbrojenia terenu, pomiary geodezyjne, obszar zurbanizowany

## Streszczenie:

Precyzyjne określenie położenia podziemnych sieci uzbrojenia terenu jest kluczowe w dziedzinie inżynierii lądowej w zakresie prac modernizacyjnych infrastruktury, podczas budowy i remontów obiektów oraz przy planowaniu i zarządzaniu przestrzenią o gęstej urbanizacji. Wiele zadań administracji publicznej takich jak: pozyskiwanie gruntów, zarządzanie własnością i planowanie zależy od wiarygodności lokalizacji uzbrojenia podziemnego. Pozwala to uniknąć zniszczeń uzbrojenia podziemnego, zminimalizować ryzyko uszkodzeń sąsiednich budynków oraz zapobiec stratom ludzkim i materialnym. Ważne jest, aby określić nie tylko lokalizację, ale również rodzaj sieci uzbrojenia podziemnego. Informacja o lokalizacji i rodzajach sieci usprawnia proces projektowania zagospodarowania terenu i wspiera zrównoważony rozwój obszarów miejskich, zwłaszcza w kontekście prac budowlanych na terenach zabudowanych i planowanych do zabudowy w dziedzinie inżynierii lądowej. Motywacją autorów do podjęcia tematu badawczego było opracowanie metodyki klasyfikacji typów sieci na podstawie bezinwazyjnie pozyskanych obrazów georadarem Leica DS2000. Autorzy zaproponowali nowy algorytm klasyfikacji bazujący na cechach geometrycznych hiperbol reprezentujących sieci podziemne. Celem pracy była również automatyzacja procesu klasyfikacji, który może wspomóc użytkownika w wyborze typu sieci na czasami bardzo zaszumionych obrazach. Echogramy pozyskano w kilkunastu różnych lokalizacjach w Otwocku i na obszarze Wojskowej Akademii Technicznej w Warszawie. Opracowany algorytm pozwala na skrócenie czasu interpretacji obrazów i selekcji obiektów podziemnych, co jest szczególnie istotne dla niedoświadczonych operatorów. Wyniki klasyfikacji wykazały, że średnia skuteczność klasyfikacji typów sieci waha się w graniach od 42% do 70% w zależności od rodzaju infrastruktury podziemnej.

Received: 2023-04-04, Revised: 2023-09-12