Weld Bead Detection Based on 3D Geometric Features and Machine Learning Approaches

PABLO RODRÍGUEZ-GONZÁLVEZ[ID][1] **AND MANUEL RODRÍGUEZ-MARTÍN**[ID][2,3]

[1]Department of Mining Technology, Topography and Structures, Universidad de León, 24401 Ponferrada, Spain
[2]Technological Department, Catholic University of Avila, 05005 Ávila, Spain
[3]TIDOP Research Group, Polytechnic School of Avila, University of Salamanca, 05003 Ávila, Spain

Corresponding author: Pablo RodrÍguez-Gonzálvez (p.rodriguez@unileon.es)

ABSTRACT Weld bead detection is essential for automated welding inspection processes. The non-invasive passive techniques, such as photogrammetry, are quickly evolving to provide a 3D point cloud with submillimeter precision and spatial resolution. However, its application in weld visual inspection has not been extensively studied. The derived 3D point clouds, despite the lack of topological information, store significant information for the weld-plaque segmentation. Although the weld bead detection is being carried out over images or based on laser profiles, its characterization by means of 3D geometrical features has not been assessed. Moreover, it is possible to combine machine learning approaches and the 3D features in order to realize the full potential of the weld bead segmentation of 3D submillimeter point clouds. In this paper, the novelty is focused on the study of 3D features on real cases to identify the most relevant ones for weld bead detection on the basis of the information gain. For this novel contribution, the influence of neighborhood size for covariance matrix computation, decision tree algorithms, and split criteria are analyzed to assess the optimal results. The classification accuracy is evaluated by the degree of agreement of the classified data by the kappa index and the area under the receiver operating characteristic (ROC) curve. The experimental results show that the proposed novel methodology performs better than 0.85 for the kappa index and better than 0.95 for ROC area.

INDEX TERMS Welding, machine learning, decision tree, weld bead, photogrammetry, 3D model, nondestructive testing.

I. INTRODUCTION

Welding is an important task in the engineering field since inadequate weld beads significantly reduce the mechanical properties of the joints, causing a reduction of their effective lifetime, and possibly their collapse with drastic consequences. For this reason, quality requirements are highly standardized, *e.g.,* [1]–[3]. However, they reach a vital point when there are involved elements subject to critical efforts, such as pressure vessels [4]. Quality control of weld joints is carried out in several phases, normally beginning with the initial visual inspection [5]. In this initial step, the external part of the weld is measured by specifically qualified (welding inspector) according to international standards to locate possible defects and external pathologies.

This first phase is carried out by expeditious methods, commonly using manual instrumentation, such as gauges kits or calipers, to generate a written report about the inspection results. The manual nature of these methods can generate inaccuracies. The certification of the weld joint is constrained by human limitations, limit of visual perception, and the precision provided by the manual instruments. Moreover, for large weld joints, there is no possibility to inspect them exhaustively, so only checked a discrete sample is checked. Additionally, not all the dimensional features can be determined precisely, for example, the three-dimensional weld misalignment [6]. As a result, the initial visual inspection phase lacks precision and reliability, with a poor performance in cost and time.

Geotechnologies encompass the set of sensors and computer algorithms that allow the acquisition, modeling and/or analysis of spatial features [7]. In spite of being associated with remote sensing, they can reach up to submillimeter scales. There are different geotechnologies available to evaluate the quality of welds and/or detect and measure surface or subsurface cracks and other defects more accurately than with the tools commonly used by the welding inspectors in

the visual inspection tasks, such as inspection gauge kits. Some of these geotechnologies include active thermography, range cameras, laser scanning [8], structured light systems [9] and photogrammetry [10]; techniques whose value to visual inspection has been demonstrated in the recent years.

The automation of noninvasive techniques by means of geotechnologies will allow the implementation of robotic platforms that track and analyze a weld joint [11], [12]. The actual approaches are either based on passive or active methods. Passive methods use optical cameras, and the bead identification is based on local thresholding over the images [13]. They are affected by illumination conditions and cannot provide any geometrical measure, and require two or more passive cameras to enable a photogrammetric reconstruction. As an alternative, the active methods, which are based on the use of an external illumination source, usually a laser beam or structured light, produce range maps of the weld joint [14]. However, the generated weld bead profiles are still inadequate for a precise bead geometric extraction [15].

One of the key issues to achieving the automation is to ensure the robust detection of the weld bead [16]. Machine learning (ML) approaches have been used in different engineering fields, and they are also beginning to be applied in welding inspection due to the geometric complexity of a weld bead [17], [18]. Neural networks have been used for prediction of weld bead geometry for automation of laser microwelding [19] and to determine the weld bead geometry, width and height, by real-time measurement from an optical camera and a laser-based sensor [20]. However, current studies are more prone to model the weld process and control of the welding parameters [21], as current, voltage and/or speed, in spite of visual inspection and maintenance tasks. A relation of ML algorithms used to model the welding process can be found in [22].

The aforementioned methods are based on spatial operations such as edge/corner detection from the images/laser profiles to identify the weld bead. The proposed methodology exploits special geometric properties of the 3D point clouds as an input of ML algorithms. The 3D features encapsulate the geometrical relationship between 3D points for a local neighborhood [23]. They can be computed from the 3D covariance matrix, or 3D structure tensor [24]. The 3D features have been used with machine learning algorithms and classification processes for complex operations such as building roof segmentation from LiDAR point clouds [25] or semantic labeling of mobile laser scanner point clouds [26], [27]. However, to the best of the authors' knowledge, no studies have been found employing 3D geometric features for ML classification of weld beads.

The aim of this work is to employ the geometric features from 3D point clouds of welds to perform automatic classification of the weld bead. The classification is based on ML decision trees to identify the key geometric features. To find the optimal weld bead detector, the number of feature attributes are analyzed according to information gain, as well as the point density for real case studies.
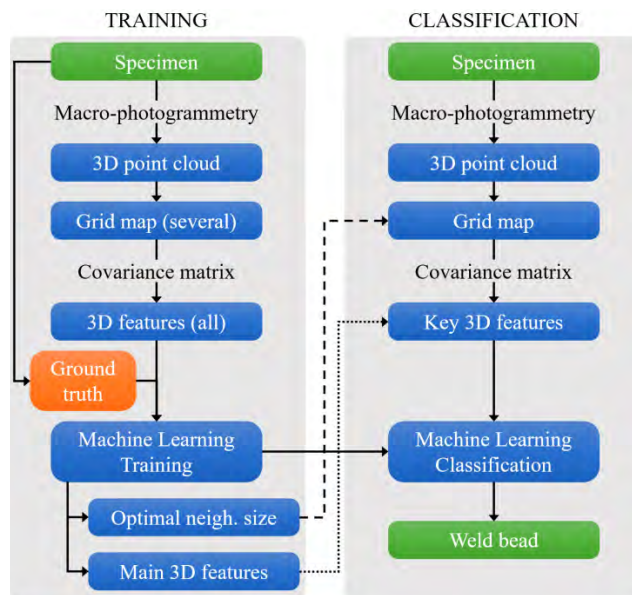

FIGURE 1. Workflow of the proposed methodology for weld bead segmentation.

The industrial application of the proposed solution would allow to undertake the complete digitalization of the welded joints and to create a digital database of them to manage their life cycle. By replacing the expeditious visual inspection of the technician with an automated and traceable process, the lack of precision, reliability and completeness of the first phase of inspection would be eliminated. In addition, the existence of a digital registry easily storable in a database of all visual inspections carried out, would allow to re-examine and / or monitor any weld joint for maintenance and prevention purposes.

The benefits in the automation of dimensional control will be especially significant for certain critical infrastructures due to the very high number of existing welded elements (e.g.: pressure vessels).

This paper is organized as follows. Section II describes the specimens employed in the study cases, the 3D geometric features, and the machine learning classifiers. The experimental results are presented and analyzed in section III. Finally, section IV explains the conclusions and future perspectives.

II. MATERIALS AND METHODS

In this section, the specimens employed for the case study and the proposed classification approach are described. The methodology is structured in three phases: the generation of the weld point cloud, the 3D feature extraction and the classification in two information classes: weld bead and plaque. The complete methodology is summed up in the following figure (Fig. 1):

As shown in the figure above, it is stablished a training step to identify the optimal parameters for machine learning classification. By means of photogrammetry (section II.2) are computed the 3D geometrical features (section II.3).

The results allow to identify two key parameters: neighborhood size (section III.1) and relevant 3D features (section III.2). The final classification results are assessed in section III.3.

A. MATERIALS

Specimens used for the study consist of welded metallic plaques (Fig. 2). The welding procedure used is tungsten inert gas welding (TIG). This kind of specimen has been chosen due to the high usability of both the material and the welding procedure in the structural and mechanical fields. The specimens present a plane butt-welding, which has been chosen due to its higher and more complex surface and curvature than other welding dispositions such as L- or T-welding.

(a)

(b)

FIGURE 2. Weld specimens. (a) Steel welding specimen with a smooth surface. (b) Steel welding specimen of low carbon steel with the rougher surface of a ship hull.

The first specimen is a low carbon steel (thickness 10 mm) which shows butt-welding with edge preparation in V (Fig. 2.a). Its dimensions are 140 mm and 100 mm in length and width, respectively. The nominal weld bead size is 20 mm. The welding procedure was a TIG used in the construction and industrial sectors. The second study case is a real naval welding specimen of low carbon steel which shows butt-welding with edge preparation in V (Fig. 2.b). Its dimensions are 300 mm in length and 7.5 mm thick. The nominal weld bead is 25 mm. The texture and roughness of both specimens are different due to their superficial peculiarities caused by the metallurgical process and surface

treatment. The first specimen presents a smooth and bright surface, while the second presents a rough surface with little brightness. The selection of two specimens with different surface finishes will allow to identify any problems in the 3D reconstruction process that can be propagated to the final weld bead classification.

For the photogrammetric reconstruction, two DSLR cameras were employed, a Canon EOS 500D and an EOS 700D, with a Sigma 50 mm macrolens (Table 1). Both cameras are commercial and semiprofessional, so their cost is more affordable than a professional DSLR.

TABLE 1. Technical specifications of the photogrammetric systems.

Camera parameters	Canon EOS 500D	Canon EOS 700D
Sensor type	CMOS	CMOS
Sensor size	22.3 x 14.9 mm (APS C)	22.3 x 14.9 mm (APS C)
Pixel size	0.0047 mm	0.0043 mm
Image size	4752 x 3168 pixels	5184 x 3456 pixels
Effective pixels	15.1 Mp	18 Mp
Lens parameters		
Principal distance	50 mm	50 mm
Field of view*	25.1º	25.1º
Aperture	f2.8 – f45	f2.8 – f45
Closer focused distance	0.188 m	0.188 m

*Field of view is not a lens parameter but results from a principal distance and sensor size. It is included for informational purposes.

B. 3D RECONSTRUCTION

As commented previously, three-dimensional reconstruction technologies are beginning to be applied in visual inspections tasks as a replacement for the inaccurate manual gadgets, such as calipers or welding gauge kits. However, the employment of geotechnologies is not widespread among welding inspectors. The main advantage of geotechnologies is the capacity to obtain 3D models in a noninvasive and remote way. These 3D models have the form of a point cloud with submillimeter spatial resolution and precision [8]. At this point, photogrammetry offers higher versatility for this task with a low cost, especially compared to metrological instruments, such as the articulated coordinate measuring machines (ACMM), at approximately 70-100 times lower cost. This manuscript is focused on the novel employment of the 3D features in weld bead detection. Please note that some geotechnologies, such as the photogrammetry records along with the 3D coordinates and the superficial radiometry of the weld cover, may allow the characterization of weld pathologies, such as corrosions.

The macro-photogrammetry procedure for 3D submillimeter reconstruction of welds was addressed in [6] and [8]. The key issues for weld bead detection lie in the image acquisition. The lighting conditions are the main constraint for the image acquisition and camera configuration. By the computation of the depth of field for small-object photogrammetry [28], it is possible to establish the optimal weld-camera

distance, ensuring that the weld bead, as well as the plaque, are correctly focused. For a predefined optimal camera shooting distance, the low-cost protocols are focused on the acquisition of several images with different angle orientations but ensuring that at any weld area is acquired from at least two different viewpoints. For a more efficient reconstruction, it is possible to arrange two or three cameras in predefined positions that can follow the weld bead longitudinal axis [29], [30].

Once the images are acquired, they are processed following the flowchart (Fig. 3) discussed in [9]: (i) tie-point extraction and matching; (ii) camera orientation; and (iii) dense reconstruction. For the practical implementation, open source software is available, such as GRAPHOS [31] or MICMAC [32].
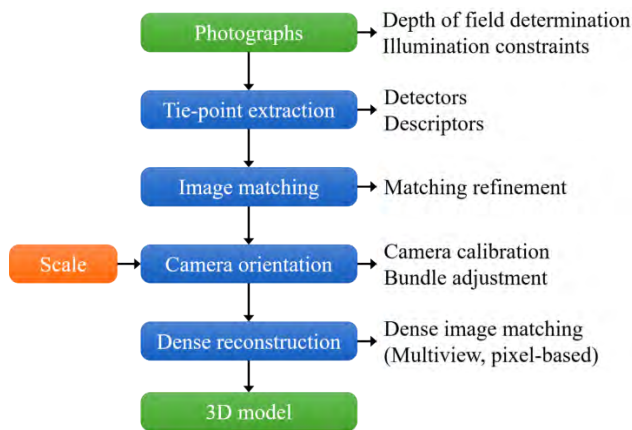

FIGURE 3. Macro-photogrammetry pipeline.

Briefly, the tie-point extraction and matching is one of the most important steps, since they constitute the framework for the orientation process (i.e. the indirect resolution of the spatial position of the images), including the camera self-calibration (i.e. the estimation of the internal camera parameters and lens distortion). Given that the constraints related to the weld, namely shape and illumination conditions, classical algorithms based on intensity levels are useless, requiring robust strategies like ASIFT [33]. Regarding camera orientation, collinearity equations are employed to determine the internal camera parameters (focal length, principal point and distortion) and exterior camera parameter (three translations and three rotations). Finally, 3D dense point cloud generation is based on the semi-global matching technique [34]. More details about the generation of dense point cloud can be found in [35].

C. 3D FEATURE EXTRACTION
Geometric features are extracted according to the depth (Z) values of the covariance matrix of a local neighborhood (square in this case). This neighborhood will be the object of study and discussed in the experimental results section since, as shown in equation (1), any element contributes to the feature computation quadratically. Instead of computing the classical geometric features as the average Z value,

or the height standard deviation, features derived from the local covariance matrix representing second-order invariant moments within the point positions are employed [36]. For the computation the initial point cloud local coordinate reference system has to be reoriented to eases the generation of a regular grid of 3D points, avoiding imbalances in the geometric features computation. As result Z values will represent the weld heights. The 3D geometric features [37]–[39] are computed on the resulting PCA analysis for local neighborhoods in a regular grip map. PCA is a procedure to obtain the three vector components of the data through the computation of eigenvectors from the covariance matrix extracted from the three-dimensional coordinate of the points. This procedure is usually employed to compute the best fitting plane, and it has already been used to compute weld misalignment [9]. Considering a point cloud of n points with coordinates x, y, z (1), where x_m, y_m, z_m are the centroid coordinates, the covariance matrix (2) for each of them is calculated from the matrix of points ($\mathbf{A}$). The covariance matrix ($\mathbf{Cov}$) has the values of the variance in the principal diagonal (2). By the diagonalization process of matrix $\mathbf{Cov}$, the eigenvectors of the covariance matrix are obtained, and as a result, the three eigenvalues (λ_1, λ_2, λ_3).

$$\mathbf{A} = \begin{pmatrix} x_1 - x_m & y_1 - y_m & z_1 - z_m \\ \vdots & \vdots & \vdots \\ x_n - x_m & y_n - y_m & z_n - z_m \end{pmatrix} \quad (1)$$

$$\mathbf{Cov} = \frac{1}{n}\mathbf{A}^t\mathbf{A} = \begin{pmatrix} \sigma_{xx} & \sigma_{yx} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \quad (2)$$

The eigenvalues correspond to the principal components of the spatial distribution of the points; they are sorted as $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and employed to compute the 3D features, which can be grouped as dimensional features [23], [36]: linearity (3), planarity (4) and sphericity (5), and other measures such as omnivariance (6), anisotropy (7), eigenentropy (9) and surface variation (10), also called the change of curvature.

$$L = \frac{\lambda_1 - \lambda_2}{\lambda_1} \quad (3)$$

$$P = \frac{\lambda_2 - \lambda_3}{\lambda_1} \quad (4)$$

$$S = \frac{\lambda_3}{\lambda_1} \quad (5)$$

$$O = \sqrt[3]{\lambda_1\lambda_2\lambda_3} \quad (6)$$

$$A = \frac{\lambda_1 - \lambda_3}{\lambda_1} \quad (7)$$

$$\Lambda = \lambda_1 + \lambda_2 + \lambda_3 \quad (8)$$

$$E = -\sum_{i=1}^{3} \frac{\lambda_i}{\Lambda} \log\left(\frac{\lambda_i}{\Lambda}\right) \quad (9)$$

$$C = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \quad (10)$$

Since it is part of the feature computation, the sum of the eigenvalues (Λ) is also considered (8).

Additionally, and as an alternative of the classical features, only the median absolute deviation – MAD of the weld heights is considered, namely, the median (m) of the absolute deviations from data median (m_Z), since it is a robust estimator of a variable dispersion [40], [41]:

$$\mathrm{MAD}_Z = m\left(|z_i - m_Z|\right) \tag{11}$$

where z_i are the weld heights obtained after a principal component analysis (PCA) reorientation prior to the regular grid generation.

A total of 9 geometric features are tested. As commented previously, the 3D features are computed for a regular square, local neighborhood.

D. MACHINE LEARNING CLASSIFICATION

Classification is implemented by modern machine learning techniques. There are two possible approaches: supervised or unsupervised classifications [42], according to whether informational classes or clusters are defined. Clusters are point groups with similar characteristics, so they do not necessarily need to correlate with classes of interest for welding inspectors and technicians.

Supervised classification was chosen to identify the more relevant 3D features, and because it provides more precise results [43]. It requires an initial training phase where a subset of points is catalogued by a human operator according to the informational classes (weld bead or plaque). The ML algorithm extracts the patterns among the points attributes and the informational classes. The generated model will be used to classify new instances. However, the models are not necessarily comparable to every weld configuration.

To choose the ML algorithm part of the proposed methodology, three different ML decision trees were tested. A complex decision was decomposed into the union of several simpler decisions [44]. In addition, among all the attributes available (in this case, the 3D geometric features), only a subset is employed for the final classification. This subset is chosen by the optimization of splitting criteria, such as the Gini index or the gain ratio.

One of their key features is that when the model is trained, the new data instances are tested against only the selected attributes according to the tree level. As a result, the computational time is improved in comparison to the classical classifiers which test all attributes for each new instance. They were chosen instead of other ML approaches, such as neural networks or support vector machines (SVM) [45], since they are not a black-box algorithm, allowing an easy interpretation of the nodes of the classification model.

The ML decision trees tested are: J48, Reptree and Hoeffding tree; all of them are implemented in the Weka data mining open-source software [46]. J48 is Weka's implementation of C4.5 release 8 [47]. This is a widely used algorithm in ML. It generates a decision tree by a recursive splitting of the dataset. The algorithm considers all possible data splits and selects the one that provides the highest ratio of information gain. The splitting criteria are based on the entropy (H),

which measures the uncertainty for the information content of a discrete variable X [48]:

$$H\left(X\right) = -\sum_{i=1}^{m} P_i \cdot \log_2 P_i \tag{12}$$

where P is the probability mass function of X, and m is all the possible values for X. Information gain (ΔH) measures the amount of information provided by the attribute. As a result, it allows the optimal split threshold that provides the maximum information to be identified. This variable is used in decision trees to establish a sequence using the attributes. Information gain measures the reduction in entropy (12), so the change in entropy when an attribute Y is used to partition X is defined as:

$$\Delta H\left(X, Y\right) = H\left(X\right) - H\left(X|Y\right) \tag{13}$$

where $H(X|Y)$ is the conditional entropy. To reduce the bias from the information gain (13), J48 employs a variant called the information gain ratio [49]. The gain ratio takes into account the number and size of branches when choosing an attribute, avoiding bias to attributes with a large number of values:

$$\textit{Gain ratio}\left(X, Y\right) = \frac{\Delta H\left(X, Y\right)}{SI\left(X, Y\right)} \tag{14}$$

where SI is the split information, which measures the potential information generated by dividing X into c classes with respect to the values of Y.

The reduced error pruning tree [50], also called the REP-Tree, is characterized by the reduction of the decision tree and the removal of tree subsections that do not improve the classification results (pruning). The splitting criteria employed is the information gain (13). This algorithm was chosen since it allows the tree depth to be controlled.

Finally, the Hoeffding tree is an incremental decision tree capable of learning from massive data streams [51]. It is based on the Hoeffding bound, which states how many instances are required at each node to carry out the classification within a predefined precision [52]. The split criterion employed in the present manuscript for the Hoeffding tree is the information gain (13).

The classification results are assessed by Cohen's kappa index from the confusion matrix and the receiving operating characteristic (ROC) curve area. The kappa statistic (κ) (15) measures the degree of agreement of categorized data [53]. It is widely used in remote sensing and ML applications to assess the classification results. It is defined in the range [−1, 1] with zero being the expected value for a random classification, 1 is a perfect agreement, and negative values indicate no agreement, but they are unlikely in practice. In remote sensing applications, the threshold for good or adequate classification is 0.85 [54]. In a more general way, according to [55], values in the range 0.81-1.00 represent an almost perfect agreement. The statistic (15) is computed on the basis of a confusion matrix with the classification results,

as shown in the following table:

$$\kappa = \frac{P_0 - P_e}{1 - P_e} \qquad (15)$$

where P_0 is an observational probability of agreement (observed accuracy) and P_e is a hypothetical expected probability of random agreement. From the confusion matrix (Table 2) can be drawn other model performance values, such as precision (16) and recall or sensitivity (17):

$$precision = \frac{TP}{TP + FP} \qquad (16)$$

$$recall = \frac{TP}{TP + FN} \qquad (17)$$

The area under the ROC curve (AUC) provides a comparison between the predicted and actual instance values in a classification by measuring the precision and the recall [56]. It measures the probability that randomly chosen instances (*e.g.,* weld bead class and plaque class) will be correctly classified. It is widely used for model comparison since it describes the model performance for a complete range of classification thresholds. ROC area is defined in the range [0, 1], with 0.5 being the expected value for a random classification. In addition, the ROC area is equivalent to the nonparametric Wilcoxon-Mann-Whitney statistic [57], which is an estimate of the probability that the classifier ranks a randomly chosen positive example higher than a negative example [58].

TABLE 2. Synthesis of confusion matrix.

		Actual classification	
		Class A (weld bead)	Class B (plaque)
Classification results (predicted)	Class A (weld bead)	True Positive (TP)	False Positive (FP)
	Class B (plaque)	False Negative (FN)	True Negative (TN)

III. EXPERIMENTAL RESULTS

First, the point cloud is, by definition, an unorganized entity, with no topology and heterogeneous density. To avoid any imbalance in the feature extraction, the input point cloud is rearranged in a grid map, consisting of regular squares of a predefined spatial resolution.

Despite photogrammetry being able to plan the output spatial resolution expressed in terms of object sample distance (OSD), the final product does not follow the grid map rigorously due to the process errors (matching, camera orientation, stereo densification) and the inherent characteristic of the scene (mainly shadows and occlusions). To generate the regular grid, the point cloud density is computed according to an ideal equilateral triangles distribution for a circular neighborhood [59]. Once the minimum spatial resolution is established, the point cloud is segmented into a regular grid

oriented according to the X and Y axis. If more than one point falls inside any grid cell, the final cell value is expressed in terms of the points' centroid. Since the point cloud axes play a role in the grid generation, prior to the segmentation, a simple PCA analysis is applied to reorient the point cloud according to the weld bead dominant direction.

The advantages of the grid generation are that it provides a 3D regular point cloud with topological information, as well as the decrease in the number of involved points in the feature extraction. This last issue allows a reduction in the computational cost and the aforementioned problems of heterogeneous point density.

The weld specimens described in section II.1 were reconstructed using macrophotogrammetry techniques. Specimen 1 was reconstrued at a very short distance, 0.38 m, due to the subtle detail of the weld bead (maximum height of 1.54 mm). A total of 23 images were acquired with a DSLR camera (Table 1), with a theoretical OSD of 36 microns. The average spatial resolution of the final 3D point cloud was 57 microns. Taking into account this value along with the a priori error of 93 microns, the regular grid map was set at 0.05 mm. The regular grip was generated by means of the open-source software, CloudCompare [60].

The naval weld (specimen 2) was reconstructed from a set of 28 images acquired at a short distance, 70 cm, with a DSLR (Table 1). Due to the larger size of this specimen, more images were taken compared to the previous case. The planned OSD was 69 microns, and the final point cloud had a spatial resolution of 71 microns. The point cloud a priori precision was verified against an ACMM by robust statistics, with an expected error of 0.204 mm. Since the spatial resolution was approximately 0.07 mm, the grid map was generated for the nearest round resolution, 100 microns. The smaller neighborhood size was 0.5 mm, 2.5 times the expected precision of the point cloud.

Second, the grid points were grouped by a square, local neighborhood to compute the representative 3D features of the clusters. They were chosen on the basis of the grid resolution. For specimen 1, they ranged between 0.50 mm and 2.50 mm, at nonregular intervals: 0.50, 0.75, 1.00, 1.25, 1.50, 2.00 and 2.50 mm. For the naval welding (specimen 2), the minimum size tested was also be 0.5 mm to provide a comparison of the higher neighborhood size of specimen 1. The rest of the sizes ranged up to 4 mm with increments of 0.5 mm; a total of 8 sizes were tested.

In Fig. 4 are shown the geometrical features computed from specimen 1 for the covariance matrices obtained for the 0.50 mm neighborhood and described in section II.3.

As a final step of the data training, a small weld section was chosen as the training area for the ML classification. For specimen 1 (Fig. 2.a) a training area of 18 x 70 mm was defined, being the 18 mm aligned according to the weld bead axis. This training area was chosen to be large enough for the computation of geometric features of the weld bead and plaque. Please note that it was chosen in an area without the presence of flaws and/or edge imperfections that could disrupt the
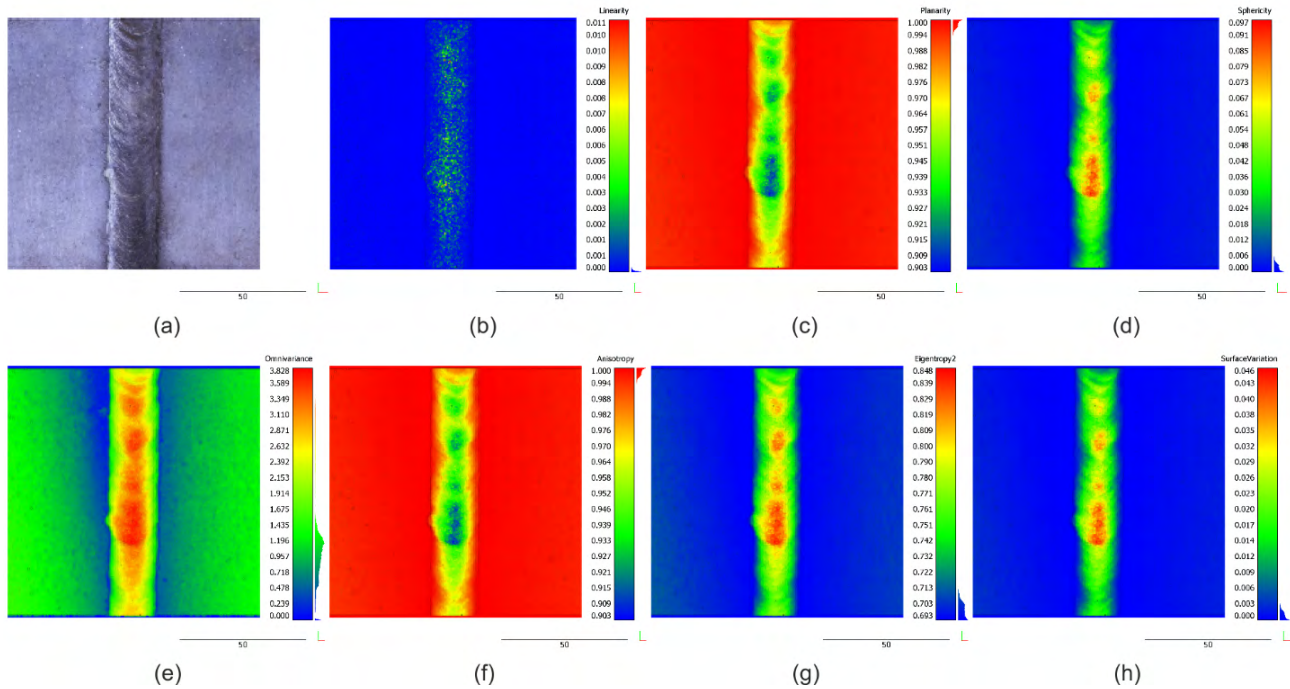
FIGURE 4. Example of 3D features: (a) original point cloud; (b) linearity, (c) planarity; (d) sphericity; (e) omnivariance; (f) anisotropy; (g) eigenentropy; (h) surface variation. The scale is in millimetres.

learning process. In this way, the proposed methodology was verified with a small training area. For specimen 2 (Fig. 2.b) the training area was similar; 20 × 75 mm.

A. LOCAL NEIGHBORHOOD SIZE ANALYSIS

To find the optimal neighborhood, different sizes were evaluated using the same training dataset. Smaller areas imply that the classifier was more sensitive to the errors propagated by the geometric features from the point cloud noise. In contrast, larger areas reduced the computational time, but they reduced the weld bead definition, since the same neighborhood contained features of both classes, so the final classification precision decreased. In Tables 3 and 4 are listed the main assessment parameters of the ML algorithms tested. It employed k-fold cross-validation, which split the training data into k equal-sized partitions [61], [62]. The classifier was trained using $k − 1$ folds, and the error was computed by checking the obtained classifier with the remaining fold. As a result, the classifier's estimated error was the average value of the errors committed in each fold. For the study case, k was set equal to 10. A total of 50 iterations were carried out for each ML algorithm and neighborhood size, for a total of 21000 tests for specimen 1 (500 for each classifier and size case). To compare the different models, a two-sided paired t-test was employed [46], using a significance level of 5%. In tables 3 and 4 are shown the results of the Kappa coefficient and weighted ROC area for the 6 ML classifier variants in relation to the 7 neighborhood areas. The J48 algorithm was set as a reference for the t-test.

TABLE 3. Kappa coefficient results for the training area of specimen 1.

| | ML Classifiers | | | | | |
| | Ref. | | Compared | | | |
Neighborhood size (mm)	J48	Reptree	Reptree (depth=6)	Reptree (depth=5)	Reptree (depth=4)	Hoeffding Tree
0.50	0.93	0.93	0.92	0.92	0.92 *	0.92
0.75	0.96	0.95	0.95	**0.96**	**0.96**	0.92 *
1.00	0.94	0.95	0.95	0.95	0.95	0.95
1.25	**0.97**	**0.96**	**0.96**	**0.96**	**0.96**	**0.96**
1.50	0.94	0.93	0.93	0.93	0.93	**0.96**
2.00	0.94	0.94	0.94	0.94	0.94	0.95
2.50	0.90	0.89	0.89	0.89	0.89	0.92

(#) Case statistically better than the reference (*) Case statistically worse than the reference. The highest values of each ML classifier are marked in bold.

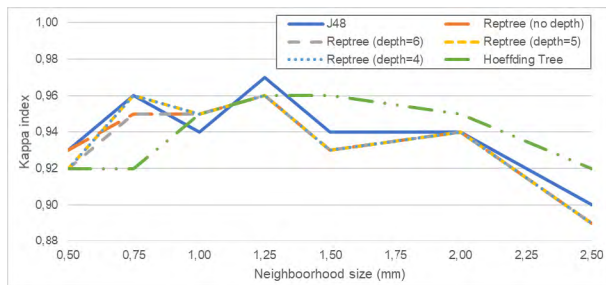**TABLE 4.** Weighted ROC area for the training area of specimen 1.

| | ML Classifiers | | | | | |
| | Ref. | | Compared | | | |
Neighborhood size (mm)	J48	Reptree	Reptree (depth=6)	Reptree (depth=5)	Reptree (depth=4)	Hoeffding Tree
0.50	0.98	0.98	0.98	0.98	0.98	0.97 *
0.75	0.98	0.98	0.98	0.98	0.98	**0.99** #
1.00	0.98	**0.99**	**0.99**	**0.99**	**0.99**	0.98
1.25	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**	0.99
1.50	0.97	0.97	0.97	0.97	0.97	**0.99**
2.00	0.97	0.96	0.96	0.96	0.96	0.98
2.50	0.96	0.94	0.94	0.94	0.94	0.98

(#) Case statistically better than the reference (*) Case statistically worse than the reference. The highest values of each ML classifier are marked in bold.

In Table 3 only two cases are statistically worse (denoted as * in the table), in relation to the reference classifier. The rest are statistically compatible at the set significance level.

In a similar way, in Table 4 all tested cases are compatible statistically, with the exception of the Hoeffding tree for the neighborhood size of 0.75 mm (denoted as # in the table).

For both parameters, the best results were achieved for neighborhood areas between 1 and 1.25 millimeters and degradation is shown as the area for the 3D feature computation was increased, but it was not statistically significant (Fig. 5). Only the Kappa index for J48 and the smaller neighborhood size was statistically worse, in relation to the rest of the sizes for the classifier (Table 5). This behavior was expected, since the number of points involved in the covariance computation was lower, and the noise from the input point cloud was beginning to appear.
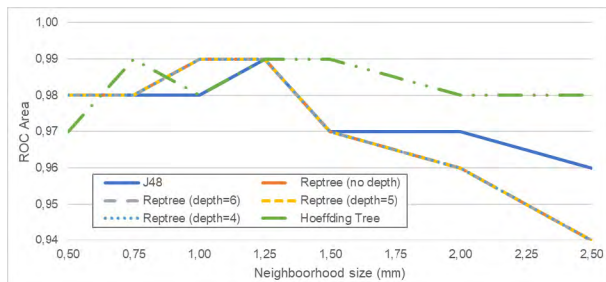
FIGURE 5. Classification assessment parameters variation according to local neighborhood size for specimen 1. (a) Kappa index. (b) ROC area.

TABLE 5. Kappa coefficient results of the J48 classifier for the training area of specimen 1.

	Neighborhood size (mm)						
	Reference		Compared				
	0.50	**0.75**	1.00	**1.25**	1.50	2.00	2.50
J48	0.93	**0.96** #	0.94	**0.97** #	0.94	0.94	0.90

(#) Case statistically better than the reference (*) Case statistically worse than the reference. The highest Kappa indexes are marked in bold.

A similar study was carried out with specimen 2, the naval welding. In spite of both specimens using butt-welding, the weld bead of the second welding was composed of several passes and had a rough surface, a higher weld height, and the plaques presented an angular misalignment. As commented previously, the training area had a size of approximately 20 × 75 mm in length and width, respectively. Tables 6 and 7 show the same experiment that Tables 3 and 4 show, namely, 50 iterations of 10-fold cross-validation for different neighborhood

TABLE 6. Kappa coefficient results for the training area of specimen 2.

	ML Classifiers					
	Ref.		Compared			
Neighborhood size (mm)	J48	Reptree	Reptree (depth=6)	Reptree (depth=5)	Reptree (depth=4)	Hoeffding Tree
0.50	0.59	0.59	0.60	0.60	0.6	0.57
1.00	0.76	0.75	0.76	0.76	0.76	0.68 *
1.50	0.78	0.79	0.79	0.79	0.79	0.71 *
2.00	0.96	**0.97**	**0.97**	**0.97**	**0.97**	0.89 *
2.50	**0.97**	0.95	0.95	0.95	0.95	0.92
3.00	**0.97**	0.92	0.92	0.92	0.92	0.91
3.50	0.86	0.9	0.90	0.90	0.90	**0.95** #
4.00	0.94	0.92	0.92	0.92	0.92	0.84

(#) Case statistically better than the reference (*) Case statistically worse than the reference. The highest values of each ML classifier are marked in bold.

TABLE 7. Weighted ROC area for the training area of specimen 2.

	ML Classifiers					
	Ref.		Compared			
Neighborhood size (mm)	J48	Reptree	Reptree (depth=6)	Reptree (depth=5)	Reptree (depth=4)	Hoeffding Tree
0.50	0.84	0.85 #	0.85	0.85	0.83	0.79 *
1.00	0.88	0.90	0.90	0.90	0.89	0.86
1.50	0.90	0.91	0.91	0.91	0.91	0.89
2.00	0.98	**0.98**	**0.98**	**0.98**	**0.98**	0.99
2.50	0.99	**0.98**	**0.98**	**0.98**	**0.98**	**1.00**
3.00	0.98	0.96	0.96	0.96	0.96	**1.00**
3.50	0.93	0.95	0.95	0.95	0.95	0.99 #
4.00	0.97	0.97	0.97	0.97	0.97	0.99

(#) Case statistically better than the reference (*) Case statistically worse than the reference. The highest values of each ML classifier are marked in bold.

sizes. Due to the wider size of the weld bead, the sizes tested ranges from 0.5 mm to 4.0 mm. A total of 24000 results were obtained.

The first conclusion drawn for the Kappa and ROC area is that smaller neighborhood sizes do not guarantee optimal classification results. The main causes of this behavior are the influence of the point cloud noise, as well as the low number of points employed for the features computations. As the grid size was set at 500 microns, only 25 points were employed in the covariance matrix computation for the smaller size tested.

The best results were between 2 and 3 millimeters, with an increase of uncertainty for larger sizes, as a result of the transition areas between the weld bead and the plaque. This tendency can be seen graphically in Fig. 6. Among the ML algorithms tested, the performance differences are not statistically significant, as in the previous study case.

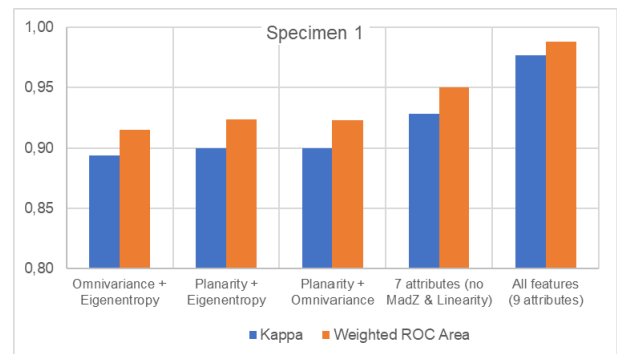### B. GEOMETRIC FEATURE CONTRIBUTION
The decision tree algorithm allows a hierarchical distribution of the features to be generated according to the ratio of gained information as a split criterion. As a result, the 3D features with a random behavior highly correlated with others were not considered. In table 8 is shown the information gain (13) and the gain ratio (14) of the geometric features for specimens 1 and 2 for a 1.25 mm and 2 mm neighborhood size, respectively. From the table it can be concluded that MAD_Z

TABLE 8. Ranked attributes. The information metric is in bits.

	Specimen 1				Specimen 2		
Information Gain		Gain Ratio		Information Gain		Gain Ratio	
Surface Variation	0.637	Planarity	0.821	Planarity	0.863	Sum(Eigenvalues)	0.928
Eigenentropy	0.637	Sum(Eigenvalues)	0.821	Sum(Eigenvalues)	0.863	Planarity	0.928
Sphericity	0.637	Anisotropy	0.720	Linearity	0.790	MAD_z	0.669
Omnivariance	0.637	Sphericity	0.720	MAD_z	0.789	Linearity	0.589
Anisotropy	0.637	Surface Variation	0.720	Surface Variation	0.631	Sphericity	0.518
Sum(Eigenvalues)	0.632	Omnivariance	0.720	Sphericity	0.613	Omnivariance	0.518
Planarity	0.632	Eigenentropy	0.720	Eigenentropy	0.613	Eigenentropy	0.518
Linearity	0.517	MAD_z	0.272	Omnivariance	0.613	Anisotropy	0.518
MAD_z	0.339	Linearity	0.245	Anisotropy	0.613	Surface Variation	0.444
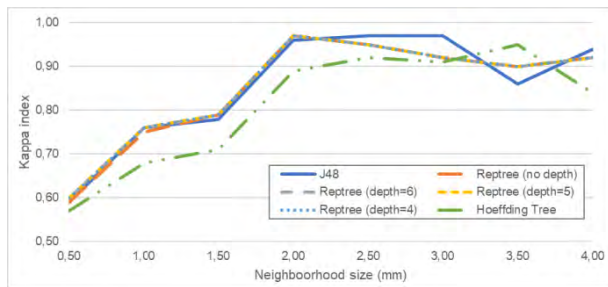
(a)
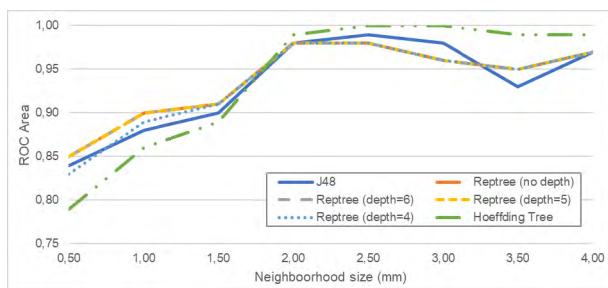
(b)

FIGURE 6. Classification assessment parameters variation according to local neighborhood size for specimen 2. (a) Kappa index. (b) ROC area.

(a)

(b)

FIGURE 7. Contribution of the geometric features to the classification. (a) Specimen 1. (b) Specimen 2.

and linearity are the features that contain less information, namely, they cause less the entropy than the other features. For specimen 2, the attributes are ranked differently, with fewer clear differences, but in both cases, the planarity is one of the main relevant attributes.

Since the application of a cut-off in Table 8 could be arbitrary, different classifications were carried out with the training dataset for the J48 algorithm and the neighborhood size of 1.25 mm. The most relevant feature combinations for specimen 1 are plotted in Fig. 7.a, where it is shown that with only 2 features, it is possible to achieve a result sufficient for the weld bead detection (Kappa index > 0.90). As expected, the planarity feature plays a critical role for the plaque detection, while the omnivariance isolates the weld bead. This combination also provides adequate classification results for the naval welding (Fig. 6.b).

C. MACHINE LEARNING CLASSIFICATION ANALYSIS

As commented in section II.4, the implementation of ML algorithms is carried out by the open source Weka data mining software [46] and the grid map generation and results visualization with CloudCompare [60]. To avoid repetition of the previous subsection, only the key classification results are plotted. Since the classification results are dependent on the neighborhood size, the predicted classes are projected over the original point cloud using a nearest neighbor interpolation with CloudCompare. Fig. 8 shows the classification results, the probability associated with the prediction, and the points tagged as classification errors on the basis of a human classification for specimen 1. The neighborhood sizes shown in Fig. 8 are 0.50 mm (a to c), 1.25 mm (d to f) and 2.00 mm (g to i). The final ML algorithm is J48, as was shown, there were no significant differences with the other ML algorithms.

For the smaller neighborhood (Fig. 8.a) there is a slight lack of definition in the weld bead borders and the presence of salt and pepper classification errors. However, since they
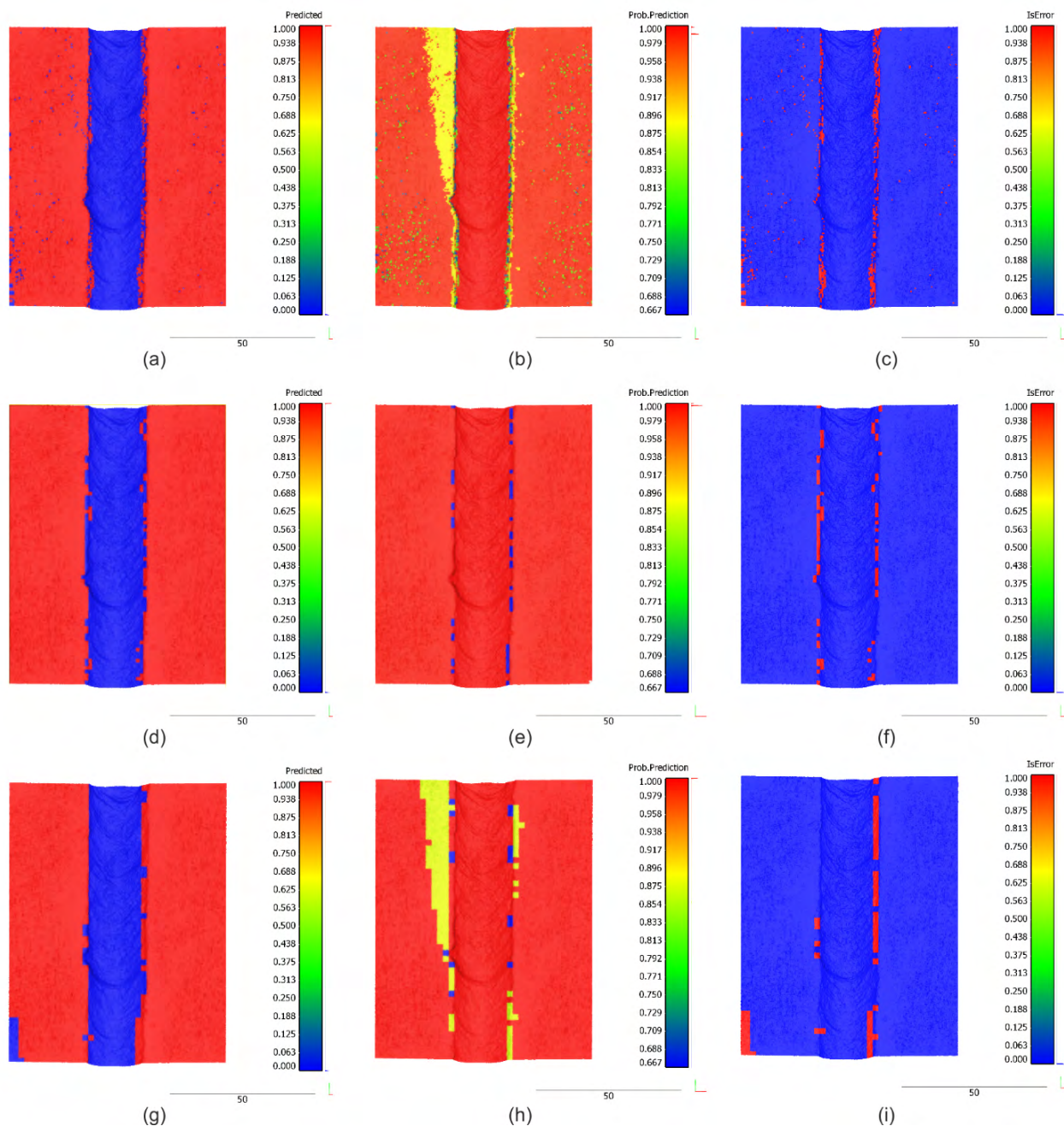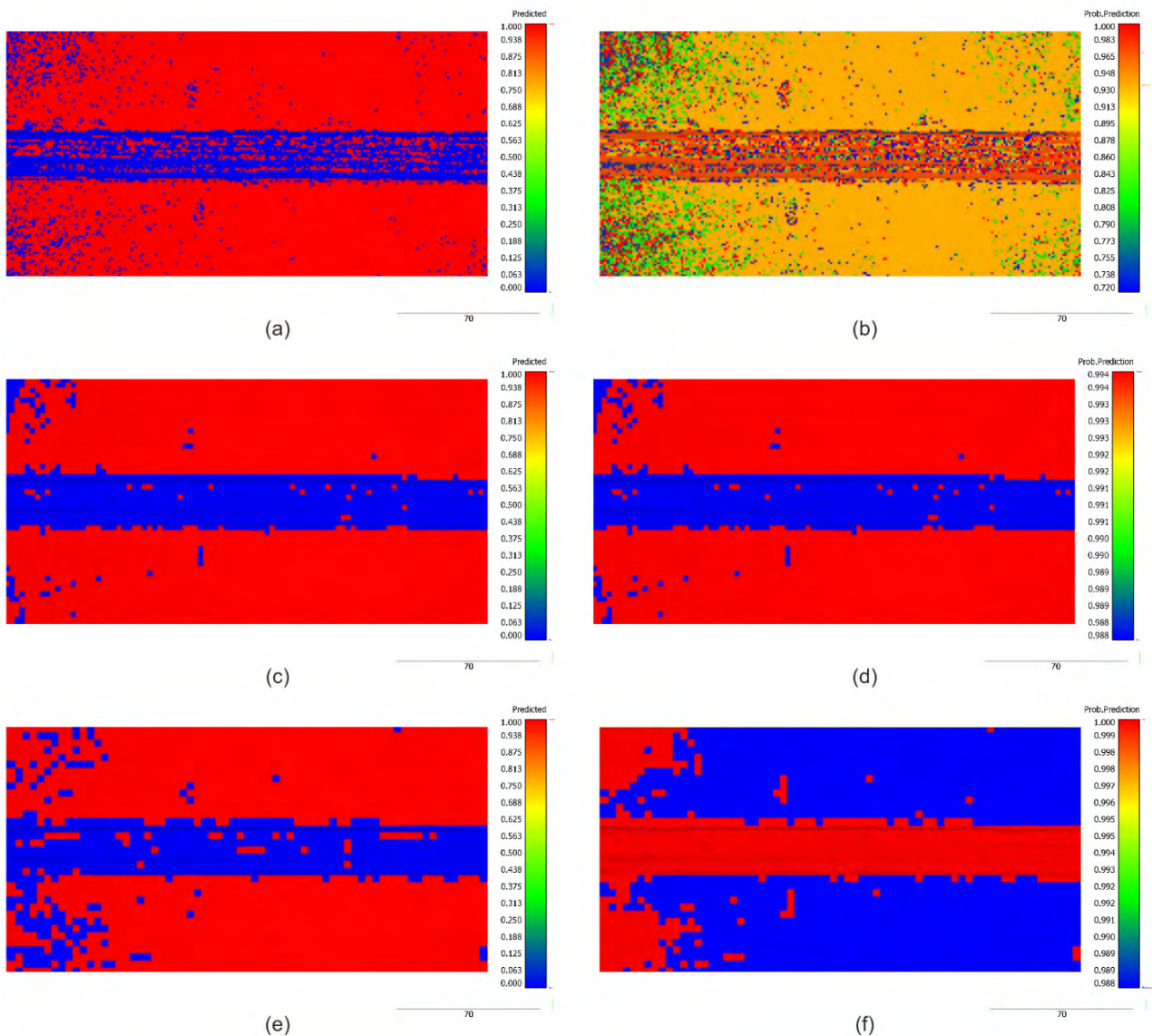
FIGURE 8. Classification results for specimen 1 for different neighborhood sizes: 0.50 mm (a), 1.25 mm (d) and 2.00 mm (g); probability associated to the prediction (b, e, h) and points labelled as classification error (c, f, i). Scale is in millimetres.

are isolated, forming small clusters, they could be removed easily by clustering analysis. Although the 0.5 mm and 1.25 mm neighborhoods yielded similar weld beads, the latter (Fig. 8.d) clearly had the best classification, both spatially, and numerically (Table 9). Fig. 8.i shows a 1 mm bias of the weld longitudinal axis (half of the 2 mm neighborhood). The results could also be accepted as adequate, since the classification performance is still good enough (Kappa index higher than 0.90), and the computational time decreased due to the low number of instances for the ML classification

which could be a clear advantage in online quality control within the automated industrial process. Please note, that in all cases, the predicted classes had a high confidence probability, approximately 85%, marked in yellow in Fig. 8. The classification results shown in Fig. 8 and the rest of the neighborhood sizes were assessed against a human classification of the weld bead. The classification assessment parameters are listed as follows:

Significant degradation of the classification results is shown for the higher size (2.50 mm), due to the excessive
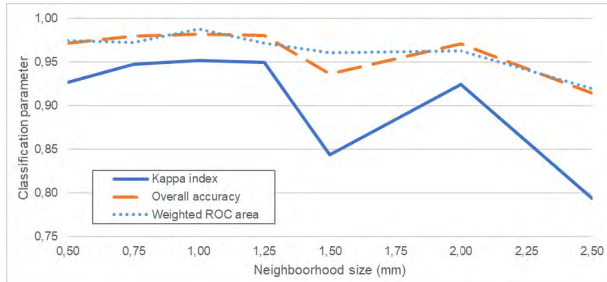
FIGURE 9. Comparison of the Kappa index, overall accuracy and weighted ROC area for the classification verification of specimen 1.

TABLE 9. Classification assessment parameters of specimen 1.

Neighborhood size (mm)	Kappa index	Overall accuracy	Weighted Precision	Weighted Recall	Weighted ROC area
0.50	0.927	0.972	0.972	0.972	0.975
0.75	0.948	0.980	0.980	0.980	0.973
1.00	**0.952**	**0.982**	**0.982**	**0.982**	**0.988**
1.25	0.950	0.981	0.981	0.981	0.972
1.50	0.844	0.937	0.942	0.937	0.961
2.00	0.925	0.971	0.971	0.971	0.963
2.50	0.794	0.915	0.924	0.951	0.920

The highest value of each parameter is marked in bold.

size in relation to the weld bead, leading to a classification uncertainty in the transition areas. Neighborhood sizes higher than 1.25 mm did not show a stable tendency (Fig. 9).

Similar to the previous study cases, the main classification assessment parameters for specimen 2 are shown in Table 10,

where the predicted classes are compared against the theoretical class to provide a vectorization by a human operator. The irregular shape of the tested naval welding constrains the result, being the maximum Kappa value reached for a neighborhood size of 2.5 mm.

In Fig. 10 are shown the classification results for the J48 decision tree of the naval welding for three different

FIGURE 10. Classification results for specimen 2 for different neighbourhood sizes: 1.00 mm (a), 2.50 mm (c) and 3.50 mm (d); probability associated with the prediction (b, d, f). Scale is in millimetres.

TABLE 10. Classification assessment parameters of specimen 2.

Neighborhood size (mm)	Kappa index	Overall accuracy	Weighted Precision	Weighted Recall	Weighted ROC area
0.50	0.539	0.854	0.851	0.854	0.842
1.00	0.712	0.908	0.907	0.908	0.866
1.50	0.633	0.857	0.897	0.857	0.884
2.00	0.739	0.903	0.927	0.903	0.924
2.50	**0.856**	**0.950**	**0.956**	**0.950**	**0.954**
3.00	0.693	0.887	0.912	0.887	0.898
3.50	0.747	0.910	0.925	0.910	0.880
4.00	0.540	0.831	0.849	0.831	0.795

The highest value of each parameter is marked in bold.

neighborhood sizes. A suboptimal one (Fig. 10.a and 10.b) where the number of points involved in the computation is not sufficient in relation to the input spatial resolution and the surface characteristics of the weld. This roughness causes the weld bead points to be classified incorrectly. This behavior was one of the aims in the specimen selection (see section II.1), where two different surface finishes were selected for evaluation. The optimal neighborhood size, 2.50 mm, is shown in Fig. 10.c and 10.d. There were still incorrectly classified weld bead points, but there were few enough to provide a robust weld detection and extraction. For this size, the transition effects begin to appear. For a high neighborhood size (3.50 mm), the weld bead limits are not sufficiently defined (Fig. 10.e).

Fig. 11 shows the aforementioned effect of neighborhood size in the covariance matrix computation. For specimen 2 the optimal value is approximately 2.5 mm (Fig. 11), where for specimen 1 the reference values start to decrease for sizes higher than 1.25 mm (Fig. 9).
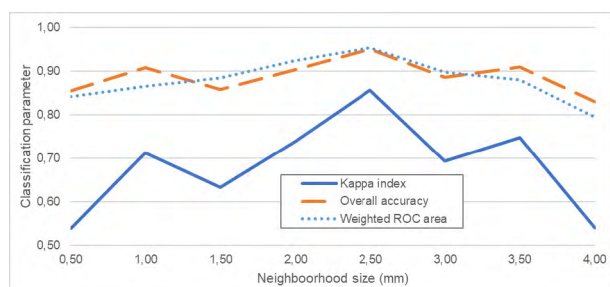

FIGURE 11. Comparison of the Kappa index, overall accuracy and weighted ROC area for the classification verification of specimen 2.

IV. CONCLUSION

In this manuscript, a novel procedure based on 3D geometric features and machine learning classifiers were tested to automatically detect the weld bead from point clouds obtained by photogrammetric techniques. In this manner, working with a 3D information approach, due to the weld complexity and weld industry requirements, opens a range of possibilities in the application of ML in welding inspection context. The geometric features were computed from the covariance matrix in an easy a quick way for a regular grip point cloud, equivalent

to a 2.5D surface. With only a subset of features, it is possible to isolate the weld bead from the plaques, the key features being identified as the planarity and the omnivariance.

Two different steel welds with different surface qualities were tested for three machine learning decision trees in six different configurations. They were analyzed against the training areas to determine whether there were any performance differences due to the split criterion employed. In all cases, the reference classifier J48 yielded no significantly different results against the reduced error pruning tree and the Hoeffding tree. All the results were assessed by the Kappa index and the area under the ROC curve.

It was determined that the neighborhood size for the feature computation varies according to the weld characteristics. However, for submillimeter point clouds, the minimum size recommended ranges from 1.00 mm. Lower sizes decrease the classification accuracy due to the low number of involved points and the presence of noise. In contrast, very wide neighborhoods do not improve the results due to the presence of impure areas, where both the weld bead and the plaque are mixed. The classification results showed high Kappa coefficients and ROC areas for both cases, approximately 0.85 - 0.95 and 0.95 - 0.99 respectively, for the optimal neighborhood size.

To conclude, the presented methodology allows the automation of the weld bead extraction for noninvasive methods for a visual inspection using the three-dimensional information provided by close-range photogrammetry technique. Given the high number of variables present in the study, further studies on automatic weld bead detection by machine learning and geometric features will focus on employing RGB values provided by photogrammetry in the ML classification to improve the precision and recall results. Moreover, some of the weld pathologies that can be detected in visual inspection can only be identified by their color, *e.g.,* corrosion pitting, so these will be tested with the addition of radiometric information for their classification. Besides deep neural network approaches will be subject of further studies to complement the propose methodology in the case of the aforementioned pathologies. Regarding the geometric pathologies, such as overlaps or undercuts (and others included in the quality standards), a higher number of classes will be considered to identify and classify them.

REFERENCES

[1] (2007). *Welding and Allied Processes–Classification of Geometric Imperfections in Metallic Materials–Part 1: Fusion Welding (ISO 6520-1:2007).* International Organization for Standardization. Accessed: Sep. 1, 2018. [Online]. Available: https://www.iso.org/standard/40229.html

[2] (2014). *Welding–Fusion-Welded Joints in Steel, Nickel, Titanium and Their Alloys (Beam Welding Excluded)–Quality Levels for Imperfections (ISO 5817:2014).* International Organization for Standardization. Accessed Sep. 1, 2018. [Online]. Available: https://www.iso.org/standard/54952.html

[3] (2018). *Welding–Arc-Welded Joints in Aluminium and Its Alloys–Quality Levels for Imperfections (ISO 10042:2018).* International Organization for Standardization. Accessed: Sep. 1, 2018. [Online]. Available: https://www.iso.org/standard/70566.html

[4] (2018). *Specification for Unfired Fusion Welded Pressure Vessels (PD 5500:2018)*. British Standards Institution. Accessed: Sep. 1, 2018. [Online]. Available: https://shop.bsigroup.com/ProductDetail/?pid= 000000000030366997

[5] (2016). *Non-Destructive Testing of Welds–Visual Testing of Fusion-Welded Joints (ISO 17637:2016)*. International Organization for Standardization. Accessed: Sep. 1, 2018. [Online]. Available: https://www.iso.org/standard/67259.html

[6] M. Rodríguez-Martín, P. Rodríguez-González, S. Lagüela, and D. González-Aguilera, "Macro-photogrammetry as a tool for the accurate measurement of three-dimensional misalignment in welding," *Automat. Construction*, vol. 71, pp. 189–197, 2016, doi: 0.1016/j.autcon.2016.08.016.

[7] P. Rodríguez-González *et al.*, "Diachronic reconstruction of lost cultural heritage sites. Study case of the medieval wall of Avila (Spain)," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLII-2, pp. 975–981, May 2018, doi: 10.5194/isprs-archives-XLII-2-975-2018.

[8] P. Rodríguez-González, M. Rodríguez-Martín, L. F. Ramos, and D. González-Aguilera, "3D reconstruction methods and quality assessment for visual inspection of welds," *Automat. Construction*, vol. 79, pp. 49–58, Jul. 2017, doi: 10.1016/j.autcon.2017.03.002.

[9] M. Rodríguez-Martín, P. Rodríguez-González, D. González-Aguilera, and J. Fernández-Hernández, "Feasibility study of a structured light system applied to welding inspection based on articulated coordinate measure machine data," *IEEE Sensors J.*, vol. 17, no. 13, pp. 4217–4224, Jul. 2017, doi: 10.1109/JSEN.2017.2700954.

[10] M. Rodríguez-Martín, S. Lagüela, D. González-Aguilera, and P. Rodríguez-González, "Procedure for quality inspection of welds based on macro-photogrammetric three-dimensional reconstruction," *Opt. Laser Technol.*, vol. 73, pp. 54–62, Oct. 2015, doi: 10.1016/j.optlastec.2015.04.011.

[11] W. Huang and R. Kovacevic, "Development of a real-time laser-based machine vision system to monitor and control welding processes," *Int. J. Adv. Manuf. Technol.*, vol. 63, pp. 235–248, Nov. 2012, doi: 10.1007/s00170-012-3902-0.

[12] J. Muhammad, H. Altun, and E. Abo-Serie, "Welding seam profiling techniques based on active vision sensing for intelligent robotic welding," *Int. J. Adv. Manuf. Technol.*, vol. 88, pp. 127–145, Jan. 2017, doi: 10.1007/s00170-016-8707-0.

[13] H. N. M. Shah, M. Sulaiman, A. Z. Shukor, Z. Kamis, and A. A. Rahman, "Butt welding joints recognition and location identification by using local thresholding," *Robot. Comput.-Integr. Manuf.*, vol. 51, pp. 181–188, Jun. 2018, doi: 10.1016/j.rcim.2017.12.007.

[14] N. R. Nayak and R. Asok, *Intelligent Seam Tracking for Robotic Welding*. London, U.K.: Springer, 2013.

[15] G. Ye, J. Guo, Z. Sun, C. Li, and S. Zhong, "Weld bead recognition using laser vision with model-based classification," *Robot. Comput.-Integr. Manuf.*, vol. 52, pp. 9–16, Aug. 2018, doi: 10.1016/j.rcim.2018.01.006.

[16] P. Kiddee, Z. Fang, and M. Tan, "An automated weld seam tracking system for thick plate using cross mark structured light," *Int. J. Adv. Manuf. Technol.*, vol. 87, pp. 3589–3603, Dec. 2016, doi: 10.1007/s00170-016-8729-7

[17] W. Chen, W. Xiong, J. Cheng, Y. Gu, and Y. Li, "Robotic vision inspection of complex joints for automatic welding," in *Proc. IEEE/ACIS 17th Int. Conf. Comput. Inf. Sci.*, Jun. 2018, pp. 246–251, doi: 10.1109/ICIS.2018.8466431.

[18] Y. Gong, Z. Lin, J. Wang, and N. Gong, "Bringing machine intelligence to welding visual inspection: Development of low-cost portable embedded device for welding quality control," *Electron. Imag.*, vol. 9, pp. 1–4, Jan. 2018, doi: 10.2352/ISSN.2470-1173.2018.09.IRIACV-279.

[19] M. I. S. Ismail, Y. Okamoto, and A. Okada, "Neural network modeling for prediction of weld bead geometry in laser microwelding," *Adv. Opt. technol.*, vol. 2013, p. 7, Oct. 2013, doi: 10.1155/2013/415837.

[20] J. F. Aviles-Viñas, R. Rios-Cabrera, and I. Lopez-Juarez, "On-line learning of welding bead geometry in industrial robots," *Int. J. Adv. Manuf. Technol.*, vol. 83, pp. 217–231, 2016, doi: 10.1007/s00170-015-7422-6.

[21] M. Leo *et al.*, "Automatic visual monitoring of welding procedure in stainless steel kegs," *Opt. Lasers Eng.*, vol. 104, pp. 220–231, 2018, doi: 10.1016/j.optlaseng.2017.09.001.

[22] S. A. Huff, "TIG welding skill extraction using a machine learning algorithm," M.S. thesis, Texas State Univ., San Marcos, TX, USA, 2017.

[23] M. Weinmann, B. Jutzi, and C. Mallet, "Feature relevance assessment for the semantic interpretation of 3D point cloud data," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. II-5/W2, pp. 313–318, Oct. 2013, doi: 10.5194/isprsannals-II-5-W2-313-2013.

[24] B. Jutzi and H. Gross, "Nearest neighbour classification on laser point clouds to gain object structures from buildings, " *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 38, pp. 1–6, 2009.

[25] Y. Xu, W. Yao, L. Hoegner, and U. Stilla, "Segmentation of building roofs from airborne LiDAR point clouds using robust voxel-based region growing," *Remote Sens. Lett.*, vol. 8, no. 11, pp. 1062–1071, 2017, doi: 10.1080/2150704X.2017.1349961.

[26] W. Yao, P. Polewski, and P. Krzystek, "Semantic labelling of ultra dense MLS point clouds in urban road corridors based on fusing CRF with shape priors," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLII-2/W7, pp. 971–976, Sep. 2017, doi: 10.5194/isprs-archives-XLII-2-W7-971-2017.

[27] D. Wang, M. Hollaus, and N. Pfeifer, "Feasibility of machine learning methods for separating wood and leaf points from terrestrial laser scanning data," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. IV-2/W4, pp. 157–164, Sep. 2017, doi: 10.5194/isprs-annals-IV-2-W4-157-2017.

[28] F. Menna, A. Rizzi, E. Nocerino, F. Remondino, and A. Gruen, "High resolution 3D modeling of the Behaim globe," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XXXIX-B5, pp. 115–120, Jul. 2012, doi: 10.5194/isprsarchives-XXXIX-B5-115-2012.

[29] H. Hu, J. Liang, Z.-Z. Xiao, Z.-Z. Tang, A. K. Asundi, and Y.-X. Wang, "A four-camera videogrammetric system for 3-D motion measurement of deformable object," *Opt. Lasers Eng.*, vol. 50, pp. 800–811, May 2012, doi: 10.1016/j.optlaseng.2011.12.011.

[30] J. Shmueli, M. A. Eder, and A. Tesauro, "A versatile stereo photogrammetry based technique for measuring fracture mode displacements in structures," *Precis. Eng.*, vol. 39, pp. 38–46, Jan. 2015, doi: 10.1016/j.precisioneng.2014.07.004.

[31] D. Gonzalez-Aguilera *et al.*, "GRAPHOS—Open-source software for photogrammetric applications," *Photogramm. Rec.*, vol. 33, pp. 11–29, Mar. 2018, doi: 10.1111/phor.12231.

[32] M. Pierrot-Deseillign and I. Clery, "Apero, an open source bundle adjustment software for automatic calibration and orientation of set of images," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XXXVIII-5/W16, pp. 269–276, Sep. 2012, doi: 10.5194/isprsarchives-XXXVIII-5-W16-269-2011.

[33] J. M. Morel and G. Yu, "ASIFT: A new framework for fully affine invariant image comparison," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 438–469, 2009, doi: 10.1137/080732730.

[34] M. Pierrot-Deseillign and N. Paparoditis, "A multiresolution and optimization-based image matching approach: An application to surface reconstruction from SPOT5-HRS stereo imagery," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XXXVI-1/W41, p. 5, 2006.

[35] F. Remondino, M. G. Spera, E. Nocerino, F. Menna, and F. Nex, "State of the art in high density image matching," *Photogramm. Rec.*, vol. 29, pp. 144–166, Jun. 2014, doi: 10.1111/phor.12063.

[36] R. Blomley, M. Weinmann, J. Leitloff, and B. Jutzi, "Shape distribution features for point cloud analysis—A geometric histogram approach on multiple scales," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. II-3, pp. 9–16, Aug. 2014, doi: 10.5194/isprsannals-II-3-9-2014.

[37] J. Demantké, C. Mallet, N. David, and B. Vallet, "Dimensionality based scale selection in 3D lidar point clouds," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XXXVIII-5/W12, pp. 97–102, Sep. 2012, doi: 10.5194/isprsarchives-XXXVIII-5-W12-97-2011.

[38] M. Weinmann, S. Urban, S. Hinz, B. Jutzi, and C. Mallet, "Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas," *Comput. Graph.*, vol. 49, pp. 47–57, Jun. 2015, doi: 10.1016/j.cag.2015.01.006.

[39] S. Guinard and L. Landrieu, "Weakly supervised segmentation-aided classification of urban scenes from 3D lidar point clouds," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLII-1/W1, pp. 151–157, May 2017, doi: 10.5194/isprs-archives-XLII-1-W1-151-2017.

[40] P. Rodríguez-González, J. Garcia-Gago, J. Gomez-Lahoz, and D. González-Aguilera, "Confronting passive and active sensors with non-Gaussian statistics," *Sensors*, vol. 14, no. 8, pp. 13759–13777, 2014, doi: 10.3390/s140813759.

[41] E. Nocerino, F. Menna, F. Remondino, I. Toschi, and P. Rodríguez-González, "Investigation of indoor and outdoor performance of two portable mobile mapping systems," in *Proc. SPIE*, vol. 10332, Jun. 2017, Art. no. 103320I, doi: 10.1117/12.2270761.

[42] J. Dougherty, R. Kohavi, and M. Sahami, "Supervised and unsupervised discretization of continuous features," in *Proc. Mach. Learn. 12th Int. Conf. Mach. Learn.*, Tahoe City, CA, USA, Jul. 1995, pp. 194–202, doi: 10.1016/B978-1-55860-377-6.50032-3.

[43] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerg. Artif. Intell. Appl. Comput. Eng.*, vol. 160, pp. 3–24, 2007.

[44] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 660–674, May/Jun. 1991, doi: 10.1109/21.97458.

[45] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh, and B. Gupta, "Deep Recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews," *J. Comput. Sci.*, vol. 27, pp. 386–393, Jul. 2018, doi: 10.1016/j.jocs.2017.11.006.

[46] (2018). *Weka 3: Data Mining Software in Java Machine Learning*. Accessed: Sep. 1, 2018. [Online]. Available: https://www.cs.waikato.ac.nz/ml/weka/

[47] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Burlington, MA, USA: Morgan Kaufmann Publishers, 1993.

[48] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.

[49] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, pp. 81–106, Mar. 1986, doi: 10.1023/a:1022643204877.

[50] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, "Random forests and decision trees," *Int. J. Comput. Sci. Issues*, vol. 9, pp. 272–278, Sep. 2012.

[51] G. Hulten, L. Spencer, and P. Domingos, "Mining time-changing data streams," in *Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, 2001, pp. 97–106, doi: 10.1145/502512.502529.

[52] P. Domingos and G. Hulten, "Mining high-speed data streams," in *Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Boston, MA, USA, 2000, pp. 71–80, doi: 10.1145/347090.347107.

[53] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960, doi: 10.1177/001316446002000104.

[54] G. M. Foody, "Harshness in image classification accuracy assessment," *Int. J. Remote Sens.*, vol. 29, pp. 3137–3158, May 2008, doi: 10.1080/01431160701442120.

[55] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977, doi: 10.2307/2529310.

[56] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognit.*, vol. 30, pp. 1145–1159, Jul. 1997, doi: 10.1016/S0031-3203(96)00142-2.

[57] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982, doi: 10.1148/radiology.143.1.7063747.

[58] C. Cortes and M. Mohri, "Confidence intervals for the area under the ROC curve," in *Proc. 17th Int. Conf. Neural Inf. Process. Syst.*, Vancouver, Canada, 2004, pp. 305–312.

[59] P. Rodríguez-González, E. Nocerino, F. Menna, S. Minto, and F. Remondino, "3D Surveying & modeling of underground passages in WWI fortifications," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XL-5/W4, pp. 17–24, Feb. 2015, doi: 10.5194/isprsarchives-XL-5-W4-17-2015.

[60] (2018). *GPL Software (Version 2.9.1)*. CloudCompare. Accessed: Sep. 1, 2018. [Online]. Available: www.danielgm.net/cc/

[61] M. Stone, "Cross-validatory choice and assessment of statistical predictions," *J. Roy. Stat. Soc. Ser. B (Methodol.)*, vol. 36, no. 2, pp. 111–147, 1974.

[62] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. 14th Int. Joint Conf. Artif. Intell.*, Montreal, Canada, vol. 2, 1995, pp. 1137–1145.

PABLO RODRÍGUEZ-GONZÁLVEZ was born in Oviedo, Spain, in 1983. He received the B.S. degree in surveying engineering from the University of Oviedo, Spain, in 2004, the M.S. degree in geodesy and cartography from the University of Salamanca, Ávila, Spain, in 2006, and the Ph.D. degree in photogrammetry and computer vision from the University of Salamanca, Spain, in 2011.

From 2013 to 2016, he was a Postdoctoral Researcher with the University of Salamanca. He has authored over 70 research articles in international journals and conference proceedings, 10 inventions, and 15 intellectual properties. His research lines include photogrammetry 3D reconstruction, accuracy assessment, and radiometric and geometric calibration of the different geomatic sensors.

He is currently an Assistant Professor with the University of León, and also the management committee member of the European COST Action CA16219.

MANUEL RODRÍGUEZ-MARTÍN was born in Ávila, Spain, in 1988. He received the B.S. degree in mechanical engineering from the University of Valladolid, Spain, in 2010, the B.S. degree in industrial engineering from Nebrija University, Spain, in 2011, the M.S. degree in mechanical engineering from the University of Salamanca, Ávila, Spain, in 2012, the Ph.D. degree from the University of Salamanca and University of Vigo, in 2015, and the L.L.B. degree from the University of Ávila, in 2017.

He has been a Research and Development Engineer, since 2012, a Researcher with the spin-off company ITOS 3D, since 2014, and held a postdoctoral position with the TIDOP Research Group, University of Salamanca, since 2015. His research line is based on geomatics sensors applied for inspection tasks in industry. He has authored over 30 research works related with this topic in international journals, books, and conference proceedings, and has worked on various research and technology transfer projects.

He is currently a Professor with the University of Ávila and University of Salamanca. He is also a Researcher with the TIDOP Research Group, University of Salamanca.

● ● ●