ONLINE FIRST September 7, 2026

Engineering Transactions, **74**, https://doi.org/10.24423/engtrans.2026.3652
Institute of Fundamental Technological Research Polish Academy of Sciences (IPPT PAN)
Université de Lorraine • Poznan University of Technology

Research Paper

Machine Learning-Based Welding Defect Recognition Using GLCM Features and k-Fold Cross-Validation: KNN and SVM Techniques

Sattar J. Kadhim[1)], AbdulBaqi AlSalait[2)], Raheem Al-Sabur[1)*], Abdel-Nasser Sharkawy[3), 4)]

[1)] *Mechanical Engineering Department, University of Basrah*
Basrah, Iraq

[2)] *Consultant of Energy Affairs, Ministry of Oil, Government of Iraq*
Baghdad, Iraq

[3)] *Mechanical Engineering Department, Faculty of Engineering, Qena University*
Qena, Egypt

[4)] *Mechanical Engineering Department, College of Engineering, Fahad Bin Sultan University*
Tabuk, Saudi Arabia

*Corresponding Author: raheem.musawel@uobasrah.edu.iq

Although radiographic inspection is one of the oldest techniques for non-destructive testing, it is still considered vital in many industrial fields to ensure the quality of welds and meet the demands of work conditions and design, as well as safety and reliability requirements. This paper presents an algorithm that identifies and categorizes welding defects in radiographic images using machine learning techniques. For this aim, two supervised classifiers are proposed and conducted: 1) K-nearest neighbor (KNN), which is a nonparametric classifier, and 2) a multiclass classifier based on a support vector machine (SVM) as a highly generalized learning-based classifier. SVM is commonly used in binary classification, but it can be adapted for multi-classification using various common methods, such as one-versus-one and one-against-all. The texture features are adopted in this paper as inputs to the classifiers, where two groups of them are used: the local binary pattern (LBP) features and the gray-level co-occurrence matrix (GLCM) extraction, to obtain the feature vector. To avoid the risk of overfitting, a 4-fold cross-validation is applied. The experimental results are reported for the two different classifiers, achieving an accuracy of 91.66 % when combining GLCM and SVM.

Keywords: welding defect, radiographic images, classification, machine learning, k-fold cross-validation.

1. Introduction

1.1. Background

To ensure the quality of welding, the traditional approach in welding environments involves using non-destructive testing (NDT) and post-weld techniques for evaluation. In welding fields, the main non-destructive tests can be categories as: visual inspection (VT) [1], ultrasonic testing (UT), radiographic testing (RT) [2], eddy current testing (ECT) [3], magnetic particle testing (MT) [4], liquid penetrant testing (PT) [5], acoustic emission testing (AE) [6], infrared thermography (IRT) [7], stereography [8], and computed tomography (CT) [9]. X-ray as a radiography test can be used on a variety of materials with varying densities, as it can penetrate a wide range of materials to detect internal faults in welds [10]. Defective regions absorb greater amounts of energy, which results in a dimmer appearance. Among the most common flaws that are possible to detect in the images of radiographic inspection are porosity (the defect that arises from the entrapment of gas), lack of fusion (the defect resulting from a lack of fusion between the base metal and the weld metal), cracks (linear discontinuities caused by the fracture of metal), solid inclusions (the defect that arises from foreign matter trapped during welding), burn-through (excessive penetration of the weld metal at the root zone of the base metal due to excessive heat), and lack of penetration (weld penetration that is less than specified, usually occurring at the root of the weld) [11–13].

1.2. Related work

The detection and classification of welding flaws are crucial in real-life and industrial environments. Furthermore, this field is considered one of the most important research topics that has been investigated. Liao and Ni [14] are considered pioneers in the use of digital radiographs to identify weld defects by comparing pixel intensities with a Gaussian pattern in the weld region. They demonstrated that similarity can be measured by the mean square error (MSE), with the lowest MSE indicating a weld line. A method was subsequently developed that focuses on the geometric features of flaws to improve the accuracy of welding defect detection using linear classifiers. It was found that having high-quality features is more important than simply having a large number of features for reliable classification [15]. Tuo *et al.* [16] achieved higher accuracy when using the Gabor filter and the gray-level co-occurrence matrix (GLCM) in welding defect classification. They applied principal component analysis (PCA) to minimize the feature vector dimensions and used K-nearest neighbors (KNN) as a comparison classifier. In the last twenty years, researchers have turned more and more to the GLCM as a practical way to study weld images [17–19]. By extracting texture features, this approach has helped make defect detection faster and more accurate, while reducing the need for manual inspection. A Gabor filter can be used to characterize texture features in an automatic system for detecting weld discontinuities, as feature extraction focuses primarily on calculating the properties of localized areas. Consequently, the GLCM displays the number of occurrences of a specific gray value in relation to another gray value in the image specific spatial context specified by the user [20]. The joint application of the Gabor filter and GLCM was later extended in welding defect classification by applying support vector machine (SVM) algorithms and evaluating performance through the receiver operating characteristic (ROC) curve [21]. This modification enabled the study of feature morphology, such as the size and shape of defects.

Artificial neural networks (ANNs) have made an influential contribution to the diagnosis of weld defects based on digital radiographic testing images, and their results have shown increasing reliability and promising validity [22]. ANNs have become increasingly popular in conjunction with other classification methods, such as KNN and SVM, to differentiate welding flaws in images from radiographic testing. Both morphological and texture features are adopted to implement the segmentation process and generate the feature vector in the classification procedure [23]. Further modifications using deep ANNs to classify the weld defects in radiographic testing images yielded accuracies of 75.83 % to 98.75 % [24]. This algorithm enables the classification of defects into more categories, including non-penetration, crack, pore, and no-defect status. Later, by applying deep learning algorithm pipelines for welding defects, the defect categories were extended to include seven flaws: porosity, undercut, crack, slag inclusion, lack of penetration, lack of fusion, and interior concavity [25].

1.3. Main contributions and outline

Based on the above discussion, further investigation into the detection and classification of weld flaws from images is warranted. Using machine learning and new advanced methods is worthy of study and evaluation. Furthermore, obtaining very high performance (accuracy, precision, and recall) for the model is the top priority.

The main contribution of this study is presented in the following points:

- an advanced algorithm that employs machine learning is proposed to detect and classify weld flaws from images with very high performance (accuracy, precision, and recall),
- the proposed algorithm involves capturing images with an ordinary digital camera, enhancing them to remove noise using a common filter, such as the adaptive median filter, and basing the classification on texture descriptors, as it has shown very high effectiveness,

– this is accomplished more specifically by utilizing the feature descriptors of the local binary pattern (LBP) and GLCM, after which the classification process commences,
– for classifiers, their selection is made using two machine learning methods: a multiclass classification based on SVM and KNN. Thus, it is possible to eliminate the need for specialized equipment or skilled workers who have experience and knowledge in the field of inspection.

The remainder of this paper is outlined as follows. Section 2 presents the proposed methodology, including the methods, the collected dataset, the classification process, and the evaluation metrics. In Sec. 3, the experimental results are presented and discussed. In Sec. 4, the paper's main points are summarized, and some potential future work is highlighted.

2. Materials and methods

2.1. Datasets preparation

Film digitization is an important step in systems used to determine welding defects. As a result, the most important factors for the performance of the above-mentioned system are the selection of an optimized scanning resolution and an acceptable level of digitization quality. Radiographic film digitization is carried out using several systems based on American Society of Mechanical Engineers (ASME) standards. The most common method is scanning, which uses light transmission (transparency adapters). The dataset images were captured with a Nikon D7000 digital camera, tuned to achieve optimal contrast. Indeed, the Nikon D7000 camera can be used for radiography using a technique called indirect digital radiography. This entails taking images of an X-ray-exposed fluorescent screen using the camera. The camera was set up at a location with a 400 mm lens, without a flash. Dataset images were prepared by converting them from RGB to grayscale and automatically cropping them. This step highlighted the welding area and a limited surrounding area. The symbols, numbers, and markings that are often associated with welding films were eliminated. Next, preprocessing models and approaches were used to minimize noise and enhance contrast, allowing the main objects in the image to be distinguished from the background, making them more discernible. Once segmentation was completed, the two feature texture extractors (LBP and GLCM) were performed to extract features that are highly distinctive for the classifiers. An automated system for identifying, recognizing, and classifying the welding defects was developed using two proposed classifiers: 1) error-correcting output coding (ECOC) with an SVM, and 2) KNN. The features and weld fault identification were obtained from these two proposed classifiers. This proposed methodology is presented in Fig. 1.

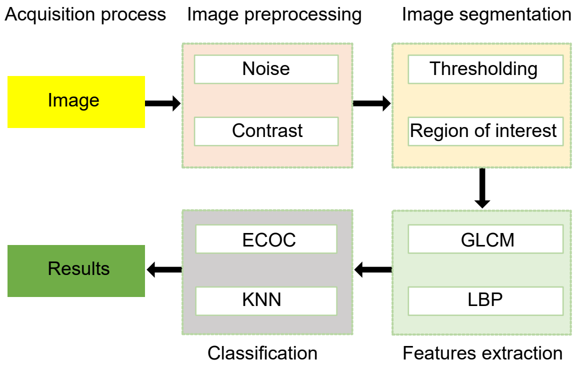

FIG. 1. Procedure for classifying welding defects (ECOC is an SVM learner).

The total number of images in the dataset is 72. The images are used to cover four types of defects, which are distributed among burn-through (BT), solid inclusions (SI), concavity (CO), and incomplete penetration (IP), as shown in Table 1. Experts in the field of radiographic interpretation established the ground truth knowledge for these images. Sample images containing BT are presented in Fig. 2.

TABLE 1. Dataset of welding defects radiographic images.

Welding defect	Number of images
SI	25
BT	17
IP	22
CO	8

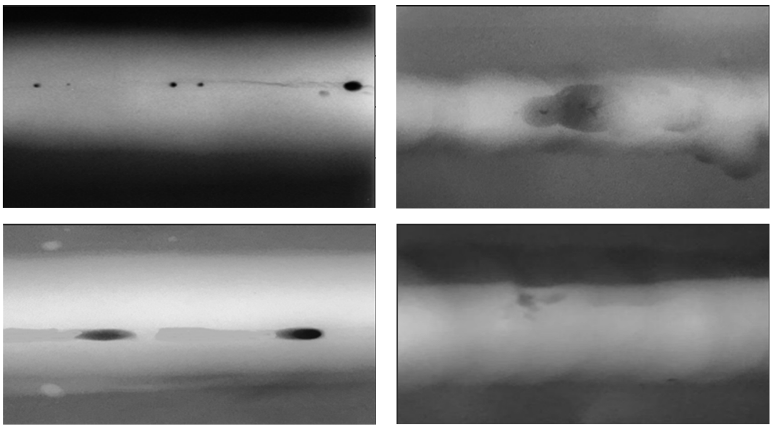

FIG. 2. Sample images containing BT. This figure contains the images from our dataset.

2.2. Image preprocessing

As an imaging system, the nature of NDT image is affected by an arbitrary signal known as noise. This signal always appears during the process of image capture, which includes coding, image transmission, and processing steps [26].

The presence of this noise could distort the primary data of the radiographic images. The source of noise in digital images can cause problems and is often considered a defect in the data acquisition process. The filtering process is a crucial and essential step in image preprocessing [27, 28]. Additionally, this process can improve image quality. The primary function of the image filtering process is to remove degradation and noise from the image, thereby achieving a high-quality image.

Image contrast enhancement refers to a type of image manipulation that modifies and improves an image appearance [29, 30]. It can be defined as the dynamic range measure, which represents the entire range of intensity values found in the image. In another definition, it is the difference between the maximum and minimum pixel values. A histogram is an essential tool to determine whether an image has a large or small dynamic range. Therefore, for high contrast, the histogram should be extended to encompass the entire dynamic range. Examples of radiographic images and image brightness improvement using different methods are presented in Fig. 3. The left side indicates the original images of IP that were contrast-stretched, then globally and next locally equalized, while the right side presents the histograms of the corresponding cases. All these improvement methods are applied to the images to find the one that provides the highest performance. The best performance is obtained by the contrast-stretching method.

2.3. Image segmentation

Segmentation can be described as dividing images into parts based on specified rules. These parts or areas can be considered homogeneous in certain visual aspects, such as color or density [31, 32]. The most effective and efficient technique for segmenting and classifying images of welding inspection films is the thresholding technique [33, 34]. An example of the segmentation process using different thresholding techniques is presented in Fig. 4. The original image (Fig. 4a) was processed using adaptive thresholding (Fig. 4b), then global thresholding (Fig. 4c), and histogram thresholding (Fig. 4d). All these segmentations methods are applied to the image. The best performance is obtained by adaptive thresholding.

2.4. GLCM

This method is a statistical approach to analyzing gray-level values by examining their spatial distribution in images, which effectively identifies distin-

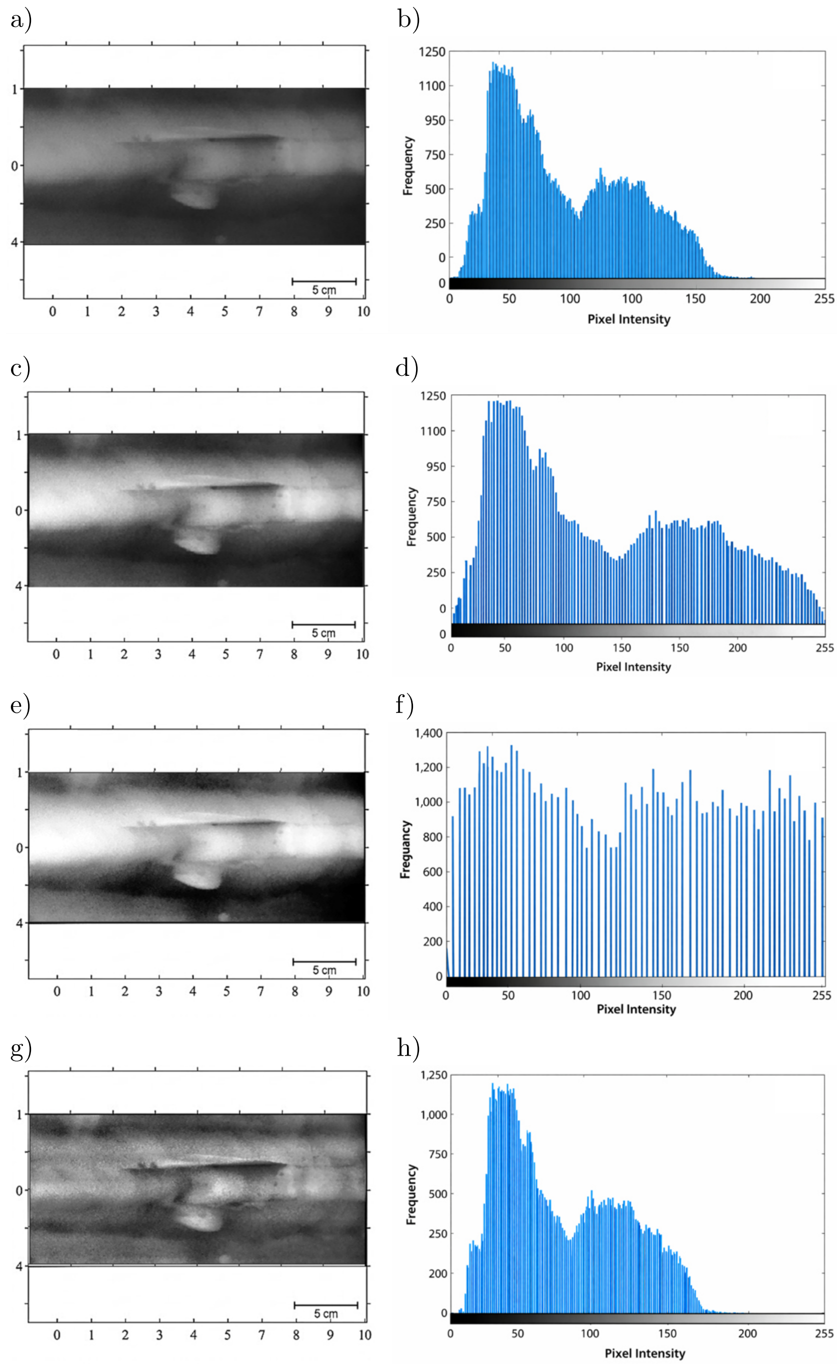

Fig. 3. Various improvement techniques carried out on an image with IP along with their corresponding histograms: a) original image, b) original image histogram, c) global equalization image, d) global histogram equalization, e) local equalization image, f) contrast limited adaptive histogram equalization (CLAHE). This figure uses the images from our dataset.

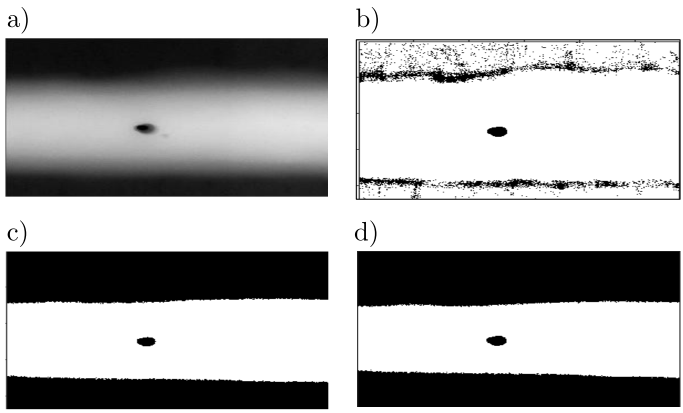

FIG. 4. Segmentation process by different thresholding techniques: a) original image, b) adaptive thresholding, c) global thresholding, d) histogram thresholding. This figure uses the images from our dataset.

guishing characteristics that can be used to find texture-related features [35–37]. The method can be summarized by creating a matrix that aggregates the total number of pixel pairs in the image. It represents the evaluation of the frequency of two gray levels within a specified region. GLCM reveals a specific relationship of several times in which a transition occurs between adjacent pixels (reflecting the correlation between two adjacent pixels within a local region). The creation of the matrices depends on both the distance and the direction. It shows that the reference pixel of intensity i has special relevance with the adjacent pixel of intensity j, which is separated by a spatial distance d and a direction θ. There are many ways for determining the spatial relationship between two neighboring pixels, considering different angles and displacements. This is clear in Fig. 5. After the matrix has been created, many statistics are derived and can be applied to the GLCM. As the number of statistics reaches 14,

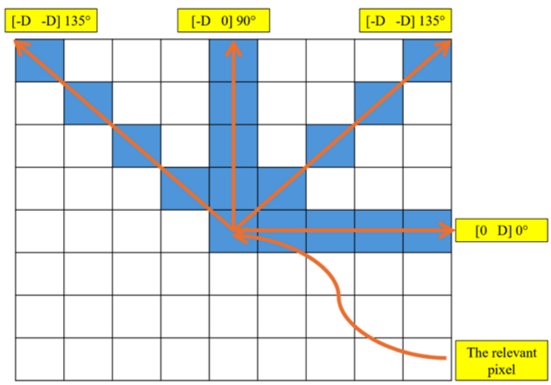

FIG. 5. Different co-occurrence matrix offsets to extract texture features.

statistical tasks are characterized by accuracy. The most widely used ones are variance, energy, homogeneity, and correlation on a large scale [38].

2.5. LBP

This factor determines the texture of an image and is based on the local differences between neighboring pixel values. In texture analysis, this factor presents a significant leap forward and is preferred overother previous techniques in various applications, as it has undergone substantial development and progress [39–41]. LBP is creative tool for texture classification [42]. The main concept of this descriptor relies on two scales: local patterns and gray-level contrast. A local binary code is extracted for each pixel found in the image by assigning the neighboring pixels with the value of the center pixel, which shows the threshold.

This process characterizes the neighborhood by multiplying the threshold values by the corresponding pixel-specific weight. Therefore, a histogram is created to accumulate the frequencies of the different binary patterns. The basic operator model relies on the eight neighboring pixels around the pixel in question [43]. The texture details of the LBP are obtained originally from a 3×3 window by the use of the gray value of the center pixel as a threshold for its neighbors. Figure 6 shows the basic LBP operator.

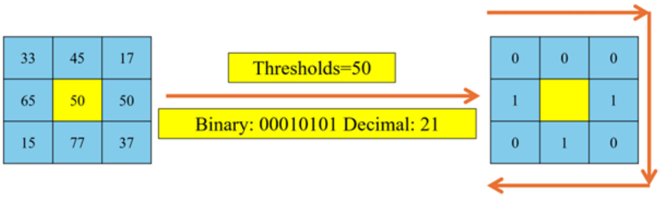

Fig. 6. Basic LBP operator.

2.6. Classification

Classification is the ability to create a model or an approach that can identify different categories in a dataset. Based on the results obtained from the training process, many outputs can be predicted. Moreover, there is a possibility of classifying these predicted data. In this study, the KNN algorithm and the SVM algorithm were used as classifiers during testing, both of which are well-suited for handling non-linear problems [44–46].

2.7. Multiclass SVM classification

Various methods have been proposed to utilize SVMs for handling multiclass problems, as they can be extended by merging multiple binary types into a multi-class classifier [47]. For present multi-class classification schemes, the one-against-all model (OAA) is the most commonly used. Classification functions are built between each category relative to the remaining categories. An M-class problem is modified into M binary classification problems. The learning step of the classifiers uses the entire training data, on the basis that the data vectors from a certain category are labelled positive and all other samples are labelled negative. In the testing step, classification is performed by calculating the margin from the linear separating hyperplane. The class with the maximum decision value is the final output. A merit of this method is that the number of binary classifiers to be constructed is equal to the number of categories. According to the maximum output value of the SVM classifiers, the label of an instance is predicted. One drawback of the OAA approach is that the dataset is significantly unbalanced after collecting the M-1 categories [48–50].

Another approach is to build a group of one-versus-one (OVO) classifiers, where the imbalance is less pronounced in the dataset than in the previous method. This is achieved by separating the M-class dataset problem into M(M-1)/2 binary classification problems and fitting a binary classification model on each. This method appears to be overly complex, as the number of classifiers obtained is often too large, making the model too intricate. Consequently, this model is not suitable for solving problems that contain a large number of categories. However, if the number of categories is small, each binary classifier is trained on a smaller subset of data. As a result, the training and tuning process of the classifier becomes more efficient and faster. To determine the class of an unknown sample, the obtained M(M-1)/2 binary classifiers are used, and the class containing the most common object samples is predicted through a voting method [48].

2.8. Evaluation metrics

To assess the quality of the classifier, a confusion matrix is used, which provides a visual representation of its predictive power. This matrix, as shown in Fig. 7 [51], includes the following classes: true positive, true negative, false positive, and false negative. These classes are defined as follows:

– true-positive (TP) means the classifier predicts positive and it is positive in fact,
– true-negative (TN) means the classifier predicts negative and it is negative in fact,
– false-positive (FP) means the classifier predicts positive and it is negative in fact,
– false-negative (FN) means the classifier predicts negative, and it is positive in fact.

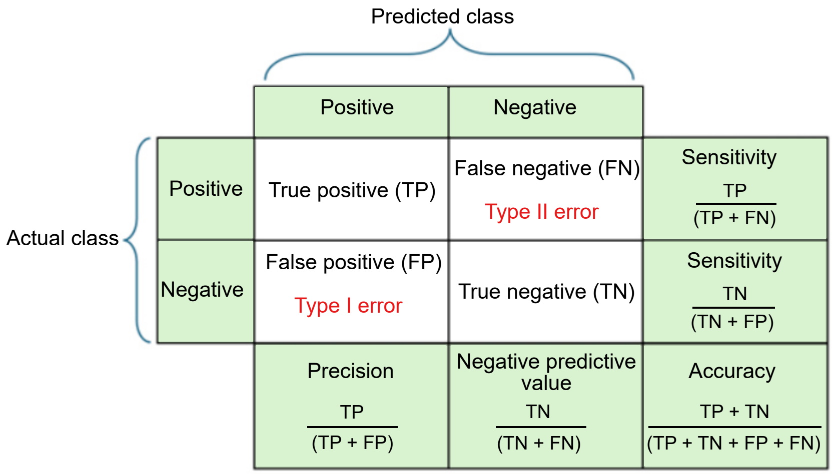

Fig. 7. Confusion matrix structure [51].

The main measures presented in the confusion matrix are discussed. Accuracy measures the closeness of the quantity to the actual/real value. It is considered a very effective metric if a relatively symmetric dataset is found [52, 53]:

$$\text{Accuracy} = \frac{\text{Correct prediction}}{\text{Total dataset}} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{2.1}$$

Measuring the percentage of correct predictions made by the proposed model is precision. It is defined by dividing the number of positive examples that are correctly classified by all the examples classified as positive by the classifier. The following form can represent it:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{2.2}$$

The third metric is recall, which indicates the percentage of all potential ground truth boxes identified. It is expressed as the number of correctly categorized positive examples divided by the total number of positive examples in the data [52]:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{2.3}$$

The experimental results from applying the proposed classifiers are presented in the next section.

3. Experimental results and discussion

Different settings for the GLCM are explored to identify the features most suitable for the classifier by conducting multiple tests that vary both distance and direction. Four directions are considered in the current work: 0°, 45°, 90°, and 135°, with the spatial pixel distance parameter set to 2. The settings for both direction and distance used for constructing GLCM are presented in Table 2.

Table 2. Setting of distance and direction for GLCM construction.

Parameter	Setting				
Item	1	2	3	4	5
Direction and angle	[0 2] (East)	[−2 2] (Northeast)	[−2 0] (North)	[−2 −2] (Northwest)	Combination of items (1 and 2)
Number of features	10	10	10	10	$10 + 10 = 20$
Features	Angular second moment, energy, entropy, contrast, maximum probability, dissimilarity, homogeneity, GLCM mean, variance and correlation				

The interpretion of all GLCM features incorporated in the current work is presented in Table 3.

Table 3. Description and visualization of all GLCM features and offests incoportated in the current work.

Offset	Distance (d) in pixels	Direction (θ)	Description
GLCM [0 1]	1	0° (horizontal)	Neighbor is 1 pixel to the right
GLCM [0 2]	2	0° (horizontal)	Neighbor is 2 pixels to the right
GLCM [−1 1]	$\sqrt{2}$	45° (diagonal)	Neighbor is 1 pixel up and 1 pixel right
GLCM [−2 2]	$2\sqrt{2}$	45° (diagonal)	Neighbor is 2 pixels up and 2 pixels right
GLCM [−1 0]	1	90° (vertical)	Neighbor is 1 pixel up
GLCM [−2 0]	2	90° (vertical)	Neighbor is 2 pixels up
GLCM [−1 −1]	$\sqrt{2}$	135° (anti-diagonal)	Neighbor is 1 pixel up and 1 pixel left
GLCM [−2 −2]	$2\sqrt{2}$	135° (anti-diagonal)	Neighbor is 2 pixels up and 2 pixels left

It should be noted that the GLCM offsets GLCM [0 2], GLCM [−2 2], GLCM [−2 0], and GLCM [−2 −2] are used for SVM and KNN. However, the GLCM offsets GLCM [0 1], GLCM [−1 1], GLCM [−1 0], and GLCM [−1 −1] are used only for SVM.

As for the parameters selected for the uniform rotation-invariant LBP, which is adopted in this investigation, are eight neighbours ($P = 8$) and their relationship to the central pixel. Therefore, the number of bins is 59, following

$P(P-1)+3$. Additionally, this method is used to achieve rotation stability by combining identical uniform patterns from different directions into a single bin. In contrast, the remaining patterns are placed in separate bins, resulting in a total of 10 distinct output bins $(P+2)$. The settings for the LBP features are presented in Table 4.

TABLE 4. Settings for the LBP features.

Parameter	Setting		
Item	1	2	3
Number of neighbors (P)	8	8	8
Rotation invariant	Yes $(P+2)$	No $(P(P-1)+3)$	Custom normalization
Result	10 features	59 features	4248 features

For the multiclass classification using SVM, the OAA and OAO approaches with a linear kernel are adopted. Concerning the KNN classifier, Euclidean distance is applied for measuring the distance between the test point and all reference points to obtain the KNN, and the number of neighbours is set as $k=1$. To select the most preferable combination of classifiers and feature extraction methods, the results are interpreted. The total accuracy, which shows the fraction of cases that are classified correctly, as well as precision and recall, are evaluated to support this selection.

Table 5 and Table 6 summarize the obtained results of the experimental evaluation, where the performance of each group of the classifier and feature extraction methods is evaluated using k-fold cross-validation as the error estimator. In the current work, $k=4$, which means that the model is trained and evaluated four times repeatedly by randomly assigning four equal-sized subgroups to satisfy the validation requirements. In each attempt, the training of the system is performed using three of the subsets, whereas the remaining subset is used for evaluation or testing.

TABLE 5. Evaluation results of classification based on multiclass SVM.

Multiclass SVM classifier	Accuracy [%]	Precision [%]	Recall [%]
GLCM $[-2\ 2]$	90.00	93.75	93.75
GLCM $[0\ 2]$	91.66	95.83	91.66
GLCM $[-2\ 0]$	90.90	93.75	95.00
GLCM $[-2\ -2]$	90.00	93.75	93.75
GLCM $\{[-2\ 2], [0\ 2]\}$	90.00	93.75	93.75
LBP 59 features	87.50	91.66	91.66
LBP 10 features	83.33	86.66	86.66

Table 6. Evaluation results of classification based on KNN.

KNN classifier	Accuracy [%]	Precision [%]	Recall [%]
GLCM [−2 2]	90.00	93.75	87.50
GLCM [0 2]	87.50	91.66	91.66
GLCM [−2 0]	90.00	93.75	87.50
GLCM [−2 −2]	90.00	93.75	87.50
GLCM {[−2 2], [0 2]}	87.50	91.66	91.66
LBP 59 features	87.50	91.66	91.66
LBP 10 features	75.00	75.00	79.17

The accuracy of the SVM classifier is higher compared with that of the KNN. Detailed results obtained by the various combinations of classifiers and features are presented in Table 7 and Table 8. Table 7 presents the value of the highest accuracy obtained when the GLCM technique is used to extract the features. Evaluating the 2-pixel neighbourhood to the east [0 2] on images leads to an accuracy value of 91.66 %. Examining the neighborhood two pixels away in all directions yields high-accuracy results. Generally, features extracted the by the use of GLCM show satisfactory results. With five tests, the achieved accuracy is 90 % or higher, and the obtained precision and recall are also 90 % or above.

Table 7. SVM multiclass classification results using GLCM features.

Multiclass SVM classifier	Accuracy [%]	Precision [%]	Recall [%]
GLCM [−2 2]	90.00	93.75	93.75
GLCM [−1 1]	83.33	92.85	91.66
GLCM [0 2]	91.66	95.83	91.66
GLCM [0 1]	88.88	93.75	87.5
GLCM [−2 −2]	90.00	93.75	93.75
GLCM [−1 −1]	84.00	90.83	93.05
GLCM [−2 0]	90.90	93.75	95.00
GLCM [−1 0]	83.33	92.85	83.3
GLCM {[−2 2], [0 2]}	90.00	93.75	93.75
GLCM {[−1 1], [0 1]}	87.50	90.66	87.5

Combining both multiclass classifiers based on SVM and LBP features (as presented in Table 8) results in high accuracy when the rotation invariance of features is not considered. Thus, it can be concluded that rotation invariance reduces accuracy, as it causes the loss of certain informative features. Therefore, the use of rotation-invariant features yield a lower accuracy of approximately 83.33 % compared with 87.50 % when omitting rotation invariance. In the lat-

TABLE 8. SVM multiclass classification results using LBP features.

Multiclass SVM classifier	Accuracy [%]	Precision [%]	Recall [%]
LBP 59 features	87.50	91.66	91.66
LBP 10 features	83.33	86.66	86.66

ter case, the classifier demonstrates strong robustness, achieving a recall and precision of 91.66 %.

Regarding the KNN classifier, several preliminary tests were conducted to determine the optimal value for the neighbourhood dimension. The conclusion reached is that the the optimal configuration is $k = 1$. However, the best results with this classifier were inferior to those obtained by the SVM classifier. The combination of the KNN classifier and GLCM features yields the best results with an accuracy of 90 %.

4. Conclusion and future work

In this paper, two machine learning approaches were proposed for detecting and classifying welding defects. These approaches are SVM and KNN. By using the dataset for training and testing strategies, the SVM showed higher performance results compared to the KNN classifier. However, the KNN clasifier relies on matching the feature profiles between the training and testing datasets. High similarity that exists among the images of welding inspectionleads to the close grouping of many extracted feature values, which is a main reason for the lower accuracy of the KNN classifier. Both GLCM and LBP techniques were used for extracting the features. When using the LBP technique, the displacement value is based on the closely adjacent neighbor (pixel). When using the GLCM technique, the displacement of two pixels adopted. Future work can be outlined as follows: 1) new software tools and classification methods can be used and assessed, 2) the current work is restricted to the spatial domain; therefore, the frequency domain can be explored, 3) investigating new and advanced supervised/unsupervised classifiers is recommended, 4) a new technique for extracting features is recommended for investigation and assessment, 5) the study can be expanded to include all types of welding defects. Moreover, expanding the dataset to include a larger number of images and a wider variety of welding defect types would improve model generalization and classification reliability.

Fundings

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

The authors declare that there are no known competing financial interests or personal relationships that could have influenced the work described in this paper.

Authors' contributions

All authors contributed equally to this paper, reviewed and approved the final manuscript.

References

1. Kumar D.D., Fang C., Zheng Y., Gao Y., Semi-supervised transfer learning-based automatic weld defect detection and visual inspection, *Engineering Structures*, **292**: 116580, 2023, https://doi.org/10.1016/j.engstruct.2023.116580.

2. Kadhim S.J., Al-Sabur R.K., Ali A.B.K., Application of different median filter algorithms for welding defects clarification in radiographic images, *University of Thi-Qar Journal for Engineering Sciences*, **11**(1): 56–61, 2020, https://doi.org/10.31663/tqujes.11.1.378(2020).

3. Wang D., Yi Q., Liu Y., Lu R., Tian G., Advanced detection and reconstruction of welding defects in irregular geometries using eddy current pulsed thermography, *NDT & E International*, **154**: 103398, 2025, https://doi.org/10.1016/j.ndteint.2025.103398.

4. Alvarado J.W.V., García L.F.C., Neira M.T., Flores J.W.V., Probability of defects detection in welded joints using the magnetic particle method, *Archives of Metallurgy and Materials*, **69**(2): 607–612, 2024, https://doi.org/10.24425/amm.2024.149789.

5. Sumardani N.I., Setiawan N.I., Nuryadin B.W., Sumardani D., Defect analysis of carbonsteel pipe welding connections using non-destructive testing with the penetrant test method, *Risenologi*, **5**(1): 38–47, 2020, https://doi.org/10.47028/j.risenologi.2020.51.72.

6. Aboali A., El-Shaib M., Sharara A., Shehadeh M., Screening for welding defects using acoustic emission technique, *Advanced Materials Research*, **1025–1026**: 7–12, 2014, https://doi.org/10.4028/www.scientific.net/amr.1025-1026.7.

7. Alasdi S.N., Al-Sabur R., In-depth thermal analysis of different PIN configurations in friction stir spot welding of similar and dissimilar alloys, *Journal of Manufacturing and Materials Processing*, **9**(6): 184, 2025, https://doi.org/10.3390/jmmp9060184.

8. Salgado-Lopez J.M., Ojeda-Elizarrarás J.L., Silva-Hernandez A., Tello-Rico J.M., Failure of an autotank repaired by welding, *Journal of Scientific and Technical Applications*, **6**(17): 1–9, http://doi.org/10.35429/jsta.2020.17.6.1.9.

9. Hamade R.F., Baydoun A.M.R., Nondestructive detection of defects in friction stir welded lap joints using computed tomography, *Materials & Design*, **162**: 10–23, 2019, https://doi.org/10.1016/j.matdes.2018.11.034.

10. Abdelkader R., Ramou N., Khorchef M., Chetih N., Bouticheb Y., Segmentation of x-ray image for welding defects detection using an improved Chan-Vese model, *Materials Today Proceedings*, **42**(5): 2963–2967, 2021, https://doi.org/10.1016/j.matpr.2020.12.806.

11. Liao T.W., Classification of weld flaws with imbalanced class data, *Expert Systems with Applications*, **35**(3): 1041–52, 2008, https://doi.org/10.1016/j.eswa.2007.08.044.

12. Chen Y., He Y., Wu L., Detection of welding defects using the YOLOV8-ELA algorithm, *Applied Sciences*, **15**(9): 5204, 2025, https://doi.org/10.3390/app15095204.

13. Orlando M., De Maddis M., Razza V., Lunetto V., Non-destructive detection and analysis of weld defects in dissimilar pulsed GMAW and FSW joints of aluminium castings and plates through 3D X-ray computed tomography, *The International Journal of Advanced Manufacturing Technology*, **132**(5–6): 2957–70, 2024, https://doi.org/10.1007/s00170-024-13576-x.

14. Liao T.W., Ni J., An automated radiographic NDT system for weld inspection: Part I – Weld extraction, *NDT & E International*, **29**(3): 157–62, 1996, https://doi.org/10.1016/0963-8695(96)00009-6.

15. Da Silva R.R., Calôba L.P., Siqueira M.H.S., Rebello J.M.A., Pattern recognition of weld defects detected by radiographic test, *NDT & E International*, **37**(6): 461–70, 2004, https://doi.org/10.1016/j.ndteint.2003.12.004.

16. Tou J.Y., Tay Y.H., Lau P.Y., Gabor filters and grey-level co-occurrence matrices in texture classification, [in:] *MMU International Symposium on Information and Communications Technologies*, pp. 197–202, 2007.

17. Valentin P., Kounalakis T., Nalpantidis L., Weld classification using gray level co-occurrence matrix and local binary patterns, [in:] *2018 IEEE International Conference on Imaging Systems and Techniques*, pp. 1–6, 2018, https://doi.org/10.1109/ist.2018.8577092.

18. Abidin Z., Anompa M.A., Muhtadan, Development of welding defects identifier application on radiographic film using gray level co-occurrence matrix and backpropagation, [in:] *AIP Conference Proceedings*, **1555**: 70–74, 2013, https://doi.org/10.1063/1.4820996.

19. Ramana E.V., Penekalapati S.V., Namala K.K., Identification of weld sub-surface defects by radiographic images using texture features, *E3S Web of Conferences*, **552**: 01017, 2024, https://doi.org/10.1051/e3sconf/202455201017.

20. Mery D., Berti M.A., Automatic detection of welding defects using texture features, *Insight – Non-Destructive Testing and Condition Monitoring*, **45**(10): 676–81, 2003, https://doi.org/10.1784/insi.45.10.676.52952.

21. Wang X., Wong B.S., Tan C.S., Recognition of welding defects in radiographic images by using support vector machine classifier, *Research Journal of Applied Sciences, Engineering and Technology*, **2**(3): 295-301, 2010.

22. Hassan J., Awan A.M., Jalil A., Welding defect detection and classification using geometric features, [in:] *2012 10th International Conference on Frontiers of Information Technology*, pp. 139–44, 2012, https://doi.org/10.1109/fit.2012.33.

23. Valavanis I., Kosmopoulos D., Multiclass defect detection and classification in weld radiographic images using geometric and texture features, *Expert Systems With Applications*, **37**(12): 7606–14, 2010, https://doi.org/10.1016/j.eswa.2010.04.082.

24. Palma-Ramírez D., Ross-Veitía B.D., Font-Ariosa P., Espinel-Hernández A., Sanchez-Roca A., Carvajal-Fals H., Nuńez-Alvarez J.R., Hernández-Herrera H., Deep convolutional neural network for weld defect classification in radiographic images, *Heliyon*, **10**(9): e30590, 2024, https://doi.org/10.1016/j.heliyon.2024.e30590.

25. Zhang W., Liu W., Yu X., Kang D., Xiong Z., Lv X., Huang S., Li Y., Deep learning-based automated detection of welding defects in pressure pipeline radiograph, *Coatings*, **15**(7): 808, 2025, https://doi.org/10.3390/coatings15070808.

26. Lone A.H., Siddiqui A.N., Noise models in digital image processing, *Global Sci-Tech*, **10**(2): 63, 2018, https://doi.org/10.5958/2455-7110.2018.00010.1.

27. Singh H., Kaur S., Sharma P., Utilizing various filtering methodologies in digital image processing, [in:] *2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT),* pp. 1–6, 2024, https://doi.org/10.1109/iceect61758.2024.10738905.

28. Al-Taie R., A review paper: digital image filtering processing, *Technium*, **3**(9): 1–11, 2021, https://doi.org/10.47577/technium.v3i9.

29. Rivera-Aguilar B.A., Cuevas E., Luque-Chang A., López J., Pérez-Cisneros M., Pixel Interaction model for contrast enhancement: bridging social science and image processing, *Applied Sciences*, **14**(23): 10887, 2024, https://doi.org/10.3390/app142310887.

30. Patel P., Bhandari A., A review on image contrast enhancement techniques, *Smart Moves Journal Ijoscience*, **5**(7): 18–22, 2019, https://scispace.com/pdf/a-review-on-image-contrast-enhancement-techniques-mn53j9lvgm.pdf.

31. Lasnel R., Frouté L., Kovscek A.R., Jolivet I.C., Creux P., Image processing and segmentation open source codes applied to FIB-SEM images of ultra-tight gas shales samples: enhanced pore space representativeness and mineral identification, *Gas Science and Engineering*, **138**: 205610, 2025, https://doi.org/10.1016/j.jgsce.2025.205610.

32. Mageswari S.U., Sridevi M., Mala C., An experimental study and analysis of different image segmentation techniques, *Procedia Engineering*, **64**: 36–45, 2013, https://doi.org/10.1016/j.proeng.2013.09.074.

33. Atta M.A., Imtiaz M., Hassan A., Saqib S., Image segmentation by using threshold techniques, *Lahore Garrison University Research Journal of Computer Science and Information Technology*, **2**(2): 1–6, 2018, https://lgurjcsit.lgu.edu.pk/index.php/lgurjcsit/issue/view/29/18.

34. Chen X., Liu C., Xie D., Miao D., Image thresholding segmentation method based on adaptive granulation and reciprocal rough entropy, *Information Sciences*, **695**: 121737, 2025, https://doi.org/10.1016/j.ins.2024.121737.

35. Haralick R.M., Shanmugam K., Dinstein I., Textural features for image classification, *IEEE Transactions on Systems Man and Cybernetics*, **SMC-3**(6): 610–621, 1973, https://doi.org/10.1109/tsmc.1973.4309314.

36. Utaminingrum F., Alqadri A.M., Somawirata I.K., Karim C., Septiarini A., Lin C.Y., Shih T.K., Feature selection of gray-level cooccurrence matrix using genetic algorithm with extreme learning machine classification for early detection of pole roads, *Results in Engineering*, **20**: 101437, 2023, https://doi.org/10.1016/j.rineng.2023.101437.

37. Iqbal N., Mumtaz R., Shafi U., Zaidi S.M.H., Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms, *PeerJ Computer Science*, **7**: 1–26, 2021, https://doi.org/10.7717/peerj-cs.536.

38. Jobanputra R., Clausi D.A., Preserving boundaries for image texture segmentation using grey level co-occurring probabilities, *Pattern Recognition*, **39**(2): 234–245, 2006, https://doi.org/10.1016/j.patcog.2005.07.010.

39. Kwak J.T., Xu S., Wood B.J., Efficient data mining for local binary pattern in texture image analysis, *Expert Systems With Applications*, **42**(9): 4529–4539, 2015, https://doi.org/10.1016/j.eswa.2015.01.055.

40. Lan S., Liao X., Fan H., Hu S., Pan Z., A multi-channel framework based Local Binary Pattern with two novel local feature descriptors for texture classification, *Digital Signal Processing*, **140**: 104124, 2023, https://doi.org/10.1016/j.dsp.2023.104124.

41. Kaur N., Nazir N., Manik, A review of local binary pattern based texture feature extraction, *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2021, https://doi.org/10.1109/ICRITO51393.2021.9596485.

42. Ojala T., Pietikäinen M., Harwood D., A comparative study of texture measures with classification based on featured distributions, *Pattern Recognition*, **29**(1): 51–59, 1996, https://doi.org/10.1016/0031-3203(95)00067-4

43. Ojala T., Pietikäinen M., Mäenpää T., Gray scale and rotation invariant texture classification with local binary patterns, [in:] *Computer Vision – ECCV 2000*, **1842**: 404–420, 2000, https://doi.org/10.1007/3-540-45054-8_27.

44. Ikram K., Djilali K., Abdennasser D., Al-Sabur R., Ahmed B., Sharkawy A.N., Comparative analysis of fouling resistance prediction in shell and tube heat exchangers using advanced machine learning techniques, *Research on Engineering Structures and Materials*, **10**(1): 253–270, 2024, https://doi.org/10.17515/resm2023.858en0816.

45. Shi B., Liu J., Nonlinear metric learning for kNN and SVMs through geometric transformations, *Neurocomputing*, **318**: 18–29, 2018, https://doi.org/10.1016/j.neucom.2018.07.074.

46. Fan Z., Xie J.K., Wang Z.Y., Liu P.C., Qu S.J., Huo L., Image classification method based on improved KNN algorithm, *Journal of Physics Conference Series*, **1930**(1): 012009, 2021, https://doi.org/10.1088/1742-6596/1930/1/012009.

47. Allwein E.L., Schapire R.E., Singer Y., Reducing multiclass to binary: a unifying approach for margin classifiers, *Journal of Machine Learning Research*, **1**: 113–141, 2000, https://www.jmlr.org/papers/volume1/allwein00a/allwein00a.pdf.

48. Liu Y., Bi J.W., Fan Z.P., A method for multi-class sentiment classification based on an improved one-vs-one (OVO) strategy and the support vector machine (SVM) algorithm, *Information Sciences*, **394–395**: 38–52, 2017, https://doi.org/10.1016/j.ins.2017.02.016.

49. Bahedh A.S., Mishra A., Al-Sabur R., Jassim A.K., Machine learning algorithms for prediction of penetration depth and geometrical analysis of weld in friction stir spot welding process, *Metallurgical Research & Technology*, **119**(3): 305, 2022, https://doi.org/10.1051/metal/2022032.

50. Al-Sabur R., Mishra A., Khalaf H.I., Mastering friction stir welding (FSW) with machine learning (ML): a comprehensive guide to algorithms and applications, [in:] *Using Computational Intelligence for Sustainable Manufacturing of Advanced Materials*, pp. 387–416, 2025, https://doi.org/10.4018/979-8-3693-7974-5.ch017.

51. Encord, *Confusion Matrix, Encord Computer Vision Glossary*, https://encord.com/glossary/confusion-matrix/ (access: 2.09.2025).

52. Sokolova M., Lapalme G., A systematic analysis of performance measures for classification tasks, *Information Processing & Management*, **45**(4): 427–37, 2009, https://doi.org/10.1016/j.ipm.2009.03.002.

53. Mahmoud K.H., Abdel-Jaber G.T., Sharkawy A.N., Neural network-based classifier for collision classification and identification for a 3-DOF industrial robot, *Automation*, **5**(1): 13–34, 2024, https://doi.org/10.3390/automation5010002.

Received September 6, 2025; revised March 5, 2026; accepted March 18, 2026;
available online March 27, 2026; version of record September 7, 2026;
published issue XXXX.