

## Article

# Defect Identification for Mild Steel in Arc Welding Using Multi-Sensor and Neighborhood Rough Set Approach

Xianping Zeng, Zhiqiang Feng \*, Xiaohong Xiang, Xin Li, Xiaohu Huang, Zufu Pan, Bingqian Li and Quan Li

Guangxi Key Laboratory of Ocean Engineering Equipment and Technology, Beibu Gulf University, Qinzhou 535100, China; 18523951688@163.com (X.Z.); bbguxxh@163.com (X.X.); lxin@bbgu.edu.cn (X.L.); 13217711127@126.com (X.H.); 7143@stu.bbgu.edu.cn (Z.P.); hhxx862@163.com (B.L.); liquan@bbgu.edu.cn (Q.L.)  
\* Correspondence: fzsqtu@163.com

**Abstract:** Welding technology plays a vital role in the manufacturing process of ships, automobiles, and aerospace vehicles because it directly impacts their operational safety and reliability. Hence, the development of an accurate system for identifying welding defects in arc welding is crucial to enhancing the quality of welding production. In this study, a defect recognition method combining the Neighborhood Rough Set (NRS) with the Dingo Optimization Algorithm Support Vector Machine (DOA-SVM) in a multisensory framework is proposed. The 195-dimensional decision-making system mentioned above was constructed to integrate multi-source information from molten pool images, welding current, and vibration signals. To optimize the system, it was further refined to a 12-dimensional decision-making setup through outlier processing and feature selection based on the Neighborhood Rough Set. Subsequently, the DOA-SVM is employed for detecting welding defects. Experimental results demonstrate a 98.98% accuracy rate in identifying welding defects using our model. Importantly, this method outperforms comparative techniques in terms of quickly and accurately identifying five common welding defects, thereby affirming its suitability for arc welding. The proposed method not only achieves high accuracy but also simplifies the model structure, enhances detection efficiency, and streamlines network training.

**Keywords:** neighborhood rough sets; defect identification; melt pool; machine learning; multi-sensor



**Citation:** Zeng, X.; Feng, Z.; Xiang, X.; Li, X.; Huang, X.; Pan, Z.; Li, B.; Li, Q. Defect Identification for Mild Steel in Arc Welding Using Multi-Sensor and Neighborhood Rough Set Approach. *Appl. Sci.* **2024**, *14*, 4978. <https://doi.org/10.3390/app14124978>

Academic Editor: Evangelos Hristoforou

Received: 18 April 2024

Revised: 28 May 2024

Accepted: 5 June 2024

Published: 7 June 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Welding, as one of the key technologies in contemporary industrial production, holds particular importance for quality monitoring in automated production processes. With the widespread use of robots in manufacturing for welding tasks, ensuring welding quality has become a crucial aspect of automated production. Although the design and optimization of process parameters before welding are critical, considering that welding itself is a highly complex process involving nonlinear interactions of multiple variables and numerous random and uncertain factors, achieving perfect surface and internal quality of welds is extremely challenging [1]. Currently, offline detection is predominantly used to assess welding quality, which significantly limits the development and efficiency improvement of automated welding technology.

Sensor technology is a key element in achieving automation and intelligence in the welding process [2]. Addressing the limitations of single sensors in terms of comprehensiveness and reliability in process information acquisition, multi-sensor information fusion leverages the complementarity of different types of information sources. This approach enables a multi-angle, multi-faceted description of the welding process and its quality characteristics, thereby facilitating the improvement of the reliability and accuracy of welding defect identification and prediction [3]. Simultaneously, the increase in the number of sensors inevitably leads to a substantial rise in the dimensionality of feature spaces, creating “big data” of the welding process. Therefore, how eliminating redundant information and noise from large-scale raw data and establishing connections between effective process

information and defect categories is of great importance for improving and perfecting online quality monitoring technology for the welding process.

Welding defect identification essentially involves classifying or predicting the target values (defect categories) of the research object (weld seam). In recent years, with the development of artificial intelligence technologies, employing machine learning or deep learning methods to develop classification models has become one of the main approaches for predicting welding defects. Currently, widely used techniques in predictive tasks include supervised learning techniques such as Artificial Neural Networks (ANN) [4], Support Vector Machines (SVM) [5], and Extreme Learning Machines (ELM) [6]. Each tool possesses distinct features. The SVM algorithm, which is based on rigorous statistical theory, can handle the segmentation problems of small sample, nonlinear datasets. However, SVM itself cannot automatically identify and eliminate irrelevant information, which, if present in excess, can slow down the system's operation and reduce classification accuracy. Additionally, key algorithm parameters such as the penalty factor  $C$  and the kernel parameter  $\gamma$  usually require empirical judgment or experimental tuning, a process that is undoubtedly time-consuming and labor-intensive. Therefore, finding a more expedient method to optimize SVM hyperparameters is particularly necessary.

In response to the processing of high-dimensional massive data generated by multi-information fusion in the welding process, this paper introduces an expanded rough set theory—Neighborhood Rough Set (NRS). One of the main contents and applications of rough set theory is feature reduction, which can eliminate redundant and non-essential features while maintaining the classification function of the system, thus forming a minimal feature set composed of key features. Therefore, using Neighborhood Rough Sets to perform feature reduction on the original dataset containing a large amount of redundant information, and then training the SVM with the streamlined dataset can effectively improve the learning efficiency of the SVM algorithm and reduce noise interference to some extent. In terms of determining the key parameters of the support vector machine, using intelligent algorithms to optimize their selection is one feasible approach to improve the classification performance of the SVM model [7]. Recently, a novel group intelligence optimization method known as the Dingo Optimization Algorithm (DOA) has been proposed. This method starts with initializing random solutions and uses fitness to assess the quality of the solutions. After multiple iterations, it finds the best overall solution, with advantages including ease of operation, good stability, and high search efficiency [8]. Therefore, combining the Dingo Optimization Algorithm with the Support Vector Machine to establish a DOA-SVM classification model and using the DOA to quickly optimize the key parameters of SVM within the definition domain can effectively improve the model's classification performance.

Despite numerous studies on welding defect detection, challenges remain in balancing accuracy and model complexity. In this work, a defect recognition method combining the NRS with the DOA-SVM in a multisensory framework is proposed to address these challenges effectively. This paper first establishes a multi-information fusion experimental platform for Gas Metal Arc Welding (GMAW), collecting and extracting features from welding process visual images, arc signals, and vibration signals, to build an information system for welding quality defect identification and prediction based on feature-level fusion (original dataset). Secondly, the Neighborhood Rough Set's feature reduction is applied to analyze the original dataset of the information system, obtaining a dataset containing fewer or minimal features. Furthermore, using the streamlined dataset to optimize the training of the DOA-SVM model, a welding defect identification and prediction model is obtained. Lastly, the model is validated with test samples, and its classification performance is compared and analyzed.

## 2. Basic Theory

### 2.1. Basic Concepts of Neighborhood Rough Sets

Classic rough set theory, which is based on the equivalence relation, can only handle discrete feature data. However, in practical applications where continuous feature data is commonly encountered, it usually requires preprocessing through discretization. There are various existing discretization methods, and different algorithms have varying impacts on the accuracy of rough set models. Nevertheless, regardless of the algorithm used, it inevitably leads to the problem of information loss [9,10]. To address this issue, Chen and Hu et al. introduce a neighborhood-based rough set model called neighborhood rough sets [11], which directly analyzes and handles continuous data.

Suppose  $(U, C \cup \{d\})$  is a decision information system that contains continuous features, where  $U$  is a non-empty finite set consisting of objects  $x_i (i = 1, 2, \dots, n)$ , called the universe;  $C$  is a collection of continuous conditional features  $a_k (k = 1, 2, \dots, m)$  describing the objects, known as the conditional feature set; and  $d$  represents discrete decision features about the objects.

Let  $B \subseteq C$  be a conditional characteristic subset in the decision information system  $(U, C \cup \{d\}), \delta > 0$ . For any  $x_i \in U$ , the neighborhood can be represented as:

$$\delta_B(x_i) = \{x_j \in U \mid D_B(x_i, x_j) \leq \delta\} \tag{1}$$

where  $D_B(x_i, x_j)$  is the distance between object  $x_i$  and  $x_j$ , with respect to  $B$  in  $U$ .

There are multiple forms of distance function  $D_B$ , among which the most representative is the Minkowski distance series [12], which includes Manhattan distance, Euclidean distance, Chebyshev distance, etc. This article adopts the following Chebyshev distance function to construct  $D_B$ , namely

$$D_B(x_i, x_j) = \max_{a_k \in B} |a_k(x_i) - a_k(x_j)| \tag{2}$$

where  $a_k(x_i)$  and  $a_k(x_j)$  are the eigenvalues of  $x_i$  and  $x_j$ , respectively, with respect to  $a_k$ .

If  $x_j \in \delta_B(x_i)$ , then for any  $a_k \in B$ , there is  $x_j \in \delta_{a_k}(x_i)$ , i.e.,

$$\delta_B(x_i) = \bigcap_{a_k \in B} \delta_{a_k}(x_i) \tag{3}$$

By the neighborhood  $\delta_B(x)$ , a neighborhood relation on the domain  $U$  of can be obtained, satisfying:

$$\begin{cases} N_B^\delta = \{(x_i, x_j) \in U^2 \mid x_j \in \delta_B(x_i)\} \\ \overline{N}_B^\delta = \bigcap_{a_k \in B} \overline{N}_{a_k}^\delta \end{cases} \tag{4}$$

The set of all neighborhood relations guided by the feature set  $C$  is denoted as  $N = \{N_B^\delta \mid B \subseteq C\}$ , and the binary tuple  $(U, N)$  is referred to as the neighborhood approximation space.

Given a decision system  $(U, C \cup \{d\}), B \subseteq C$ , if  $Y \in U/\{d\}$  is divided by the decision attribute  $d$  into decision classes, then in the neighborhood approximation space  $(U, N)$ , the lower and upper approximation sets of  $Y$  concerning  $B$  and  $\delta$  are, respectively.

$$\begin{cases} \underline{N}_B^\delta(Y) = \{x \in U \mid \delta_B(x) \subseteq Y\} \\ \overline{N}_B^\delta(Y) = \{x \in U \mid \delta_B(x) \cap Y \neq \emptyset\} \end{cases} \tag{5}$$

The relative positive domains of  $d$  in  $(U, N)$  with respect to  $B$  and  $\delta$  are

$$POS_B^\delta(d) = \bigcup_{Y \in U/\{d\}} \underline{N}_B^\delta(Y) \tag{6}$$

Feature selection (attribute reduction) is a core aspect of rough set theory and methodology research. Its objective is to select and construct a set containing the minimum or fewer features while maintaining the classification ability of decision information systems. This

set is referred to as the minimal feature set or minimal reduction, to eliminate redundant features and improve classifier performance.

The classification ability of a decision information system can be measured by approximate classification quality (also called approximate dependence). In the neighborhood approximation space  $(U, N)$ , the approximate classification qualities of  $d$  with respect to  $B$  and  $\delta$  are:

$$\gamma_B^\delta(d) = \text{Card}(POS_B^\delta(d)) / \text{Card}(U) \tag{7}$$

where  $\text{Card}(\cdot)$  is the cardinality of the set.

Let  $B \subset C$ , from Formula (3), for any  $a \in C - B$ , The following equation is obtained

$$\delta_{B \cup \{a\}}(x) = \delta_B(x) \cap \delta_a(x) \subseteq \delta_B(x) \tag{8}$$

Based on Equations (5)–(7), it follows that

$$POS_{B \cup \{a\}}^\delta(d) = \bigcup_{Y \in U / \{d\}} N_{B \cup \{a\}}^\delta(Y) \supseteq \bigcup_{Y \in U / \{d\}} N_B^\delta(Y) = POS_B^\delta(d) \tag{9}$$

$$\gamma_{B \cup \{a\}}^\delta(d) = \text{Card}(POS_{B \cup \{a\}}^\delta(d)) / \text{Card}(U) \geq \text{Card}(POS_B^\delta(d)) / \text{Card}(U) = \gamma_B^\delta(d) \tag{10}$$

Equations (8)–(10) indicate that the neighborhood decreases as the number of features increases, while the relative positive region and approximate classification quality increase with an increasing number of features.

By utilizing the monotonicity property of approximate classification quality, a forward-type feature selection algorithm can be constructed, also known as the one-by-one addition method. This method involves selecting one feature at a time from the candidate subset and adding it to the minimal feature set until there is no further change in its approximate classification quality.

### 2.2. Support Vector Machine Classification Principles

SVM, a significant classification algorithm in the traditional machine learning field, was initially proposed by Chervonenkis in 1963 [13], However, the commonly seen form of SVM nowadays (i.e., soft-margin) was introduced by Cortes et al. in 1993 and published in 1995 [14]. SVM has been regarded as the most successful and high-performing algorithm in machine learning until deep learning emerged as a dominant force in 2012. By incorporating VC-dimensionality theory and the principle of structural risk minimization, SVM aims to find an optimal classification hyperplane that minimizes structural risk. Compared to traditional machine learning algorithms, SVM exhibits exceptional performance and finds broad applications. The fundamental concept behind SVM is to identify the best hyperplane for performing classification tasks. This hyperplane not only accurately separates two sets of samples (even when empirical risk is zero), but also maximizes the margin between classes. Consequently, SVM demonstrates superior robustness and generalization capabilities while maintaining high accuracy levels for classification tasks. The working principle of a support vector machine can be understood as follows: given training samples, it constructs a decision surface represented by a hyperplane that maximizes separation between positive and negative examples [15]. Figure 1 illustrates a schematic diagram depicting support vector machine classification.

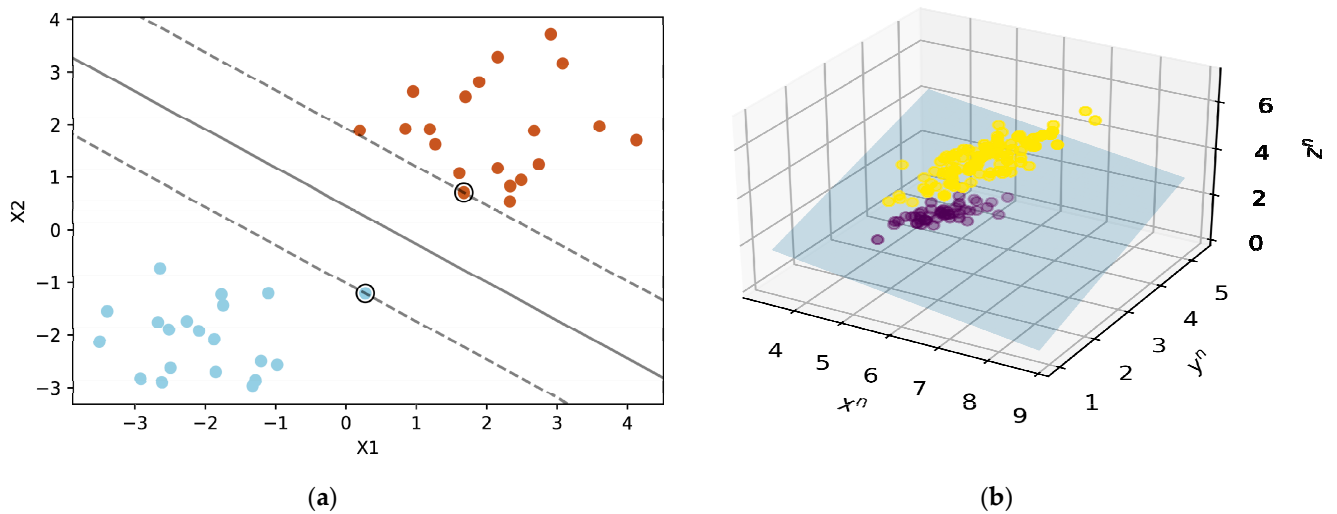
The general model of SVM is described as follows [16]: Given the sample set  $(x_i, y_i)$ ,  $i = 1, 2, \dots, l$ .  $x \in R^l, y_i \in \{+1, -1\}$ , the classification decision function constructed by SVM is defined as follows:

$$f(x) = \text{sgn}\{(w \cdot x + b)\} = \text{sgn}\left(\sum_{i=1}^l \alpha_i y_i K(x_i, x) + b^*\right) \tag{11}$$

Here,  $K(x_i, x)$  is a symmetric kernel function satisfying the Mercer condition. The task of constructing a nonlinear categorical decision function becomes solving the following optimization problem:

$$\begin{cases} \max W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \text{s.t. } \sum_{i=1}^l \alpha_i y_i = 0 \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, l \end{cases}$$

The vector corresponding to the coefficients  $\alpha_i \neq 0$  is known as the support vector, and  $C$  is a constant that reflects the degree of penalty for misclassified samples. SVM training is solving a convex quadratic programming problem with boundary constraints and linear equation constraints.



**Figure 1.** Schematic diagram of support vector machine classification. (a) two-dimensional space; (b) higher dimensional space.

### 3. Support Vector Machine Based on Neighborhood Rough Set Data Preprocessing and Optimization of Dingo Algorithm

Aiming at the difficulties of SVM methods in dealing with high-dimensional large-scale data and the sensitivity to anomalous samples, we propose a DOA-SVM classification method based on the preprocessing of neighborhood rough set data. This method preprocesses the training dataset in two main ways. On the one hand, we use the outlier detection algorithm based on isolated forests [17] proposed by Liu et al. in 2008 to eliminate those outlier samples or noisy data that are mixed with other categories of samples. On the other hand, we use neighborhood rough sets for attribute simplification based on attribute importance. Finally, the support vector machine is optimized with the dingo optimization algorithm to build a support vector machine model (DOA-SVM).

#### 3.1. Attribute Reduction for Neighborhood Rough Sets

Feature importance is a crucial basis for feature selection. In forward-type algorithms, the importance of a candidate feature relative to a subset of features depends on the change in approximate classification quality caused by adding that feature. Let  $B \subset C$ , for any  $a \in C - B$ , the importance of  $a$  relative to  $B$  in the neighborhood approximation space  $(U, N)$  be expressed as

$$\gamma_{B \cup \{a\}}^\delta(d) = \text{Card}(\text{POS}_{B \cup \{a\}}^\delta(d)) / \text{Card}(U) \geq \text{Card}(\text{POS}_B^\delta(d)) / \text{Card}(U) = \gamma_B^\delta(d) \quad (12)$$

The objective of neighborhood rough set feature selection is to find an attribute subset that has the same classification capability as the original data while excluding redundant attributes. Although there are usually multiple reductions for a given decision table, finding one of them is sufficient in most applications.

To ensure that the importance of smaller value features is not overshadowed by larger value features during the classification process, we normalize each feature. Based on attribute importance indicators, we construct a greedy algorithm for feature selection. The basic idea of this algorithm is shown in Figure 2.

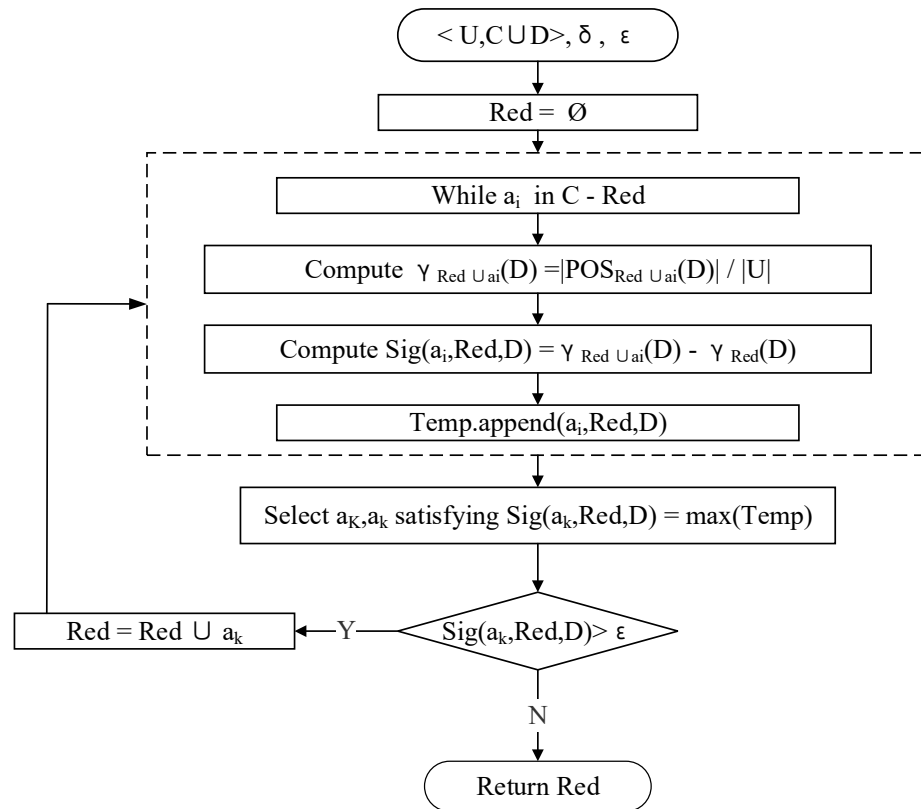


Figure 2. Attribute reduction based on a forward greedy search algorithm.

The reduction process is roughly described: the reduction set starts with the empty set, each time adding an attribute to the reduction set and calculating the attribute importance of the temporary reduction set after adding the attribute  $a_i$ . From these temporary reduction sets, the set with the greatest attribute importance is selected. Judge whether the maximum attribute importance  $sig(red \cup a_k)$  is greater than the importance threshold  $\epsilon$  set by us. If it is, add the attribute  $a_k$  to the collection  $red$  and update it; if it is less than or equal to  $\epsilon$ , exit the program.

### 3.2. Support Vector Machines Based on the Dingo Optimization Algorithm

DOA (Dingo Optimization Algorithm) [18] is a novel intelligent optimization algorithm proposed by Herna'n Peraza-Va'zquez et al. in 2021 based on the social behavior of Australian dingoes. DOA has been widely used due to its strong optimality-seeking ability and fast convergence rate, etc. DOA is inspired by the hunting strategies of dingoes, which include three strategies, namely pack attack, persecution, and scavenging, and introduces an optimality-seeking individual survival rate rule.

The basic working principle of DOA is as follows:

- (1) Initialization: a group of Australian dingoes is randomly generated, with each dingo representing a solution.
- (2) The fitness value of each dingo is calculated based on the fitness function.

- (3) Strategy selection: each dingo judges which strategy to choose based on a comparison of random numbers and probability values.

Strategy 1: Group Attack, which mimics the behavior performed by large creatures when they are found, is described by the formula:

$$\vec{x}_i(t+1) = \beta_1 \sum_{k=1}^{na} \frac{[\vec{\varphi}_k(t) - \vec{x}_i(t)]}{na} - \vec{x}_*(t) \tag{13}$$

where  $\vec{x}_i(t+1)$  is the adjusted position of the individual (particle or agent) in the next iteration,  $na$  is in the interval  $[-2, n/2]$  (integer value) chosen randomly.  $\vec{\varphi}_k(t)$  is the best position found by the  $k$ -th individual (particle or agent) with  $\varphi \leq X$ ,  $X$  is the dingoes' population randomly generated.  $\vec{x}_i(t)$  is the current position of the individual, and  $\vec{x}_*(t)$  is the individual's best historical position found,  $\beta_1$  is in the interval  $[-2, 2]$ , and is a random number in the interval; it is a weight that affects the individual's ability to explore.

Strategy 2: Persecution, which simulates the behavior of small creatures such as rabbits, pheasants, etc. that are found in hot pursuit until they are captured, is described as:

$$\vec{x}_i(t+1) = \vec{x}_*(t) + \beta_1 \cdot e^{\beta_2} \cdot (\vec{x}_{r_1}(t) - \vec{x}_i(t)) \tag{14}$$

Moreover,  $\beta_2$  is also a random number in the interval  $[-1, 1]$ ,  $r_1$  is a unique individual of the current best position found in the interval,  $\vec{x}_i(t)$  is the position of the individual, and  $i \neq r_1$ .

Strategy 3: Scavenger, which simulates the behavior of a dingo that walks randomly in its habitat to find a carrion and hunt it. The formula is described as:

$$\vec{x}_i(t+1) = \frac{1}{2} [e^{\beta_2} \cdot \vec{x}_{r_1}(t) - (-1)^\sigma \cdot \vec{x}_i(t)] \tag{15}$$

Here,  $\sigma$  represents a randomly generated 0 or 1.

- (4) Survival judgments are re-renewed for individuals with less than 30% survival, which is calculated using the formula:

$$\text{survival}(i) = \frac{\text{fitness}_{\text{max}} - \text{fitness}(i)}{\text{fitness}_{\text{max}} - \text{fitness}_{\text{min}}} \tag{16}$$

The position update formula is:

$$\vec{x}_i(t) = \vec{x}_*(t) + \frac{1}{2} [\vec{x}_{r_1}(t) - (-1)^\sigma * \vec{x}_{r_2}(t)] \tag{17}$$

fitness represents the fitness of the objective function; survival is the survival rate,  $r_1$  and  $r_2$  are random different values, and  $\vec{x}_{r_2}(t) \neq \vec{x}_{r_1}(t)$ .

DOA has been widely applied to a number of optimization problems and engineering problems, including functional optimization, combinatorial optimization, machine learning, and continuous engineering problems [8,19,20].

The penalty factor  $C$  and the kernel parameter  $\gamma$  in the Support Vector Machine (SVM) model have a large impact on the computational accuracy, and the selection of appropriate  $C$  and  $\gamma$  parameters has a great effect on the computational results of the model. In this section, according to the characteristics of the dingo algorithm, which offers dynamic search capabilities and avoids settling for local optimal solutions easily, the  $C$  and  $\gamma$  parameters of SVM are optimized, and the specific optimization steps are as follows. The DOA-SVM Flowchart is shown in Figure 3:

- (1) Generate a random initial population based on the upper and lower limits of the  $C$ ,  $\gamma$  parameters.

- (2) SVM cross-validation using the generated individual location parameters (C, gamma) to obtain the fitness value for each individual.
- (3) Make a judgment about which strategy to choose for each individual.
- (4) Update the position of each individual and its fitness value.
- (5) Calculate the individual survival rate and update the position of the individual with a low survival rate based on the best position.
- (6) Judge whether the DOA algorithm satisfies the stopping criterion; if not, return to step 3; if it does, terminate the optimization search, output the current most position (C, gamma), build the DOA-SVM model, and end the procedure.

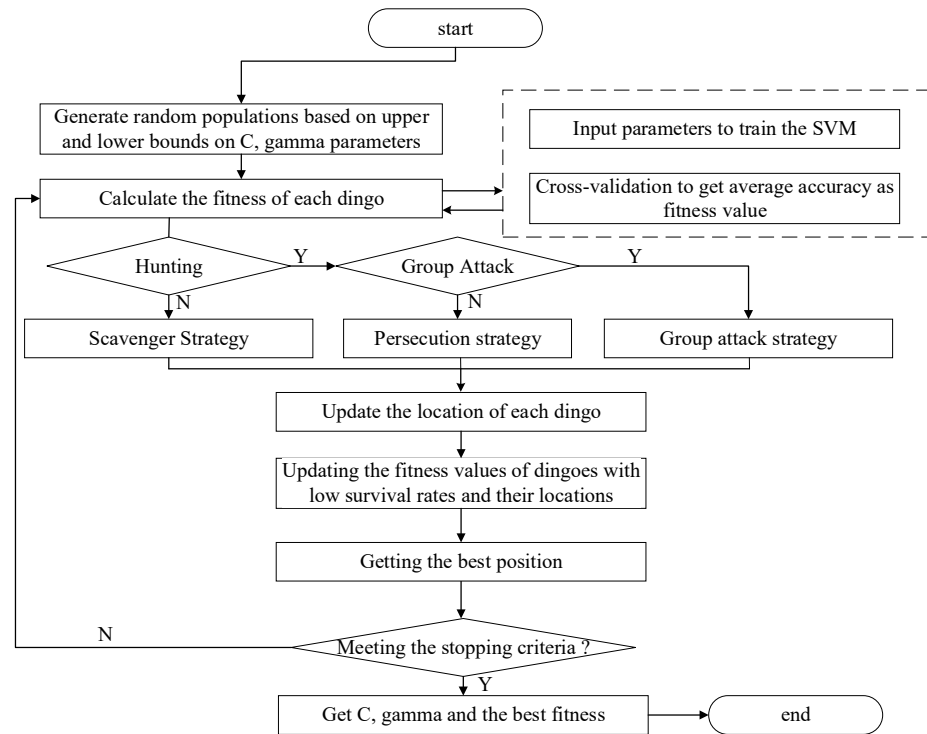


Figure 3. DOA-SVM flowchart.

### 3.3. Experimental Analyses

In order to verify the validity of the method in this paper, the feature approximation of the selected four UCI datasets is performed using the rough set (RS) and neighborhood rough set (NRS) with equal-width discretization, and the details of the data descriptions are shown in Table 1; the popular K-Nearest Neighbors (KNN), Classification and Regression Tree (CART), Support Vector Machine (SVM), and Dingo Optimization Algorithm Support Vector Machine (DOA-SVM) are used for the feature reduction in the UCI datasets. and CART, SVM, and DOA-SVM classification algorithms to evaluate the classification ability of the selected features through 5-fold cross-validation. Evaluate the classification ability of the selected features.

Table 1. UCI Dataset Information Table.

Data Set	Sample Size	Number of Attributes	Number of Categories
Wine	178	14	3
Iono	351	34	2
Sonar	208	60	2
Glass	214	9	7

When calculating the sample neighborhood, the data are first normalized to reduce the impact of the results because of the different attribute scales. Then, the outliers are

removed to reduce the interference of the anomalous data; at the same time, the data are normalized concerning the literature [11] for attribute reduction concluded that the neighborhood threshold takes a value size between [0.2, 0.4] is more effective. In this paper, after a 5-fold cross-validation comparison, the best result is obtained when  $\delta = 0.2$ , the combined results are the best. Table 2 records the number of attributes corresponding to the attribute reduction of the dataset by classical rough set and neighborhood rough set. Table 3 records the classification performance ability of the selected features in each classification algorithm; RS represents the features after reduction based on equal-width discretization and rough sets, NRS represents the features after approximation based on neighborhood rough sets, and Raw represents the original features without approximation.

**Table 2.** Comparison of Attribute Reduction.

Data Set	Attribute Index	
	RS	NRS
Wine	[13, 4, 10, 3, 1, 12, 8, 11, 2]	[13, 10, 6, 7, 12]
Iono	[1, 5, 6, 12, 32, 29, 8, 33, 7, 20, 17, 33, 34]	[1, 5, 19, 30, 16, 31, 6, 3]
Sonar	[54, 19, 45, 35, 27, 23, 20, 29, 24, 16]	[44, 12, 21, 31, 24, 1]
Glass	[7, 3, 8, 5, 2, 4, 1, 9]	[8, 4, 7, 5, 9, 2, 3, 1, 6]

**Table 3.** Comparison of classification algorithm accuracy.

		Classification Accuracy (%)			
		CART	KNN	SVM	DOA-SVM
Wine	RS	90.33	94.41	94.01	95.07
	NRS	91.33	96.08	96.48	97.32
	Raw	90.11	94.93	97.18	97.89
Iono	RS	85.6	84.32	86.33	87.18
	NRS	90.51	86.32	91.03	92.06
	Raw	87.71	85.18	89.69	90.96
Sonar	RS	70.62	69.67	75.18	84.33
	NRS	69.48	72.04	71.68	74.88
	Raw	71.7	78.8	75.3	75.9
Glass	RS	59.91	55.60	59.64	61.7
	NRS	66.00	63.08	66.61	69.59
	Raw	68.00	63.08	67.33	69.01
Average Value		78.44	78.63	80.87	82.99

From Table 2, it can be observed that both the equal-width discretization combined with the rough set and the neighborhood rough set-based simplification method can effectively reduce the data dimensionality. It is worth noting that compared with the rough set approach, the neighborhood rough set on the “Sonar” dataset has a reduction rate as high as 90%, which clearly indicates that the neighborhood rough set is a more effective means of data dimensionality reduction and redundancy data elimination.

Table 3 demonstrates the classification accuracies of each dataset feature under four different classification algorithms. A comparison of the accuracy of rough sets and neighborhood rough sets shows that the latter is usually better than the former. The underlying reason for this difference is that the choice of discretization method has a non-negligible impact on the data, which may lead to a reduction in the relevance of the original data and loss of some information. However, neighborhood rough sets successfully overcome this drawback. By comparing the classification accuracy of Neighborhood Rough Set (NRS) with that of the original data (Raw), it can be found that the accuracy of NRS is equal to or higher than that of the original features, which indicates that NRS can effectively pick out the representative data features, and thus eliminate the redundant features. Finally, in terms of the average value of classification accuracy, DOA-SVM reaches 82.99%, which is

significantly better than the other three classification algorithms, and this result highlights the superiority of the DOA-SVM model proposed in this paper.

#### 4. Welding Multi-Source Sensing Systems and Their Information Processing and Knowledge Modeling

##### 4.1. Multi-Source Information Sensing and Acquisition

In this paper, for the original robotic welding system, a set of intelligent sensing systems was designed and developed, integrating vibration sensors and molten pool image acquisition, which were mounted on the welding workpiece and the robot end torch, respectively. This system equips the welding production robot with sensing organs similar to eyes and skin, enabling it to sense the surrounding environment in real time. The intelligent sensing system uses high-performance high-speed camera CCDs, vibration sensors, and current Hall sensors to collect real-time images of the welding dynamic molten pool, vibration signals, and current parameters. Through the welding local end controller, multi-source heterogeneous data during the welding process can be collected, displayed, and stored for a short period in real time. In this way, the multi-source sensing system for the arc welding robot can be enhanced to more effectively control the quality, stability, and reliability of the welding process. Figure 4 shows the schematic structure of this intelligent sensing system.



**Figure 4.** Multi-source sensing system.

This study aims to explore the effects of different GMAW parameters on the welding quality of low-carbon steel and to achieve real-time accurate identification of welding defects using multi-sensor information fusion technology. The experimental specimens were low-carbon steel plates with dimensions of 300 mm × 150 mm × 6 mm, configured as butt joints in the flat welding position, without using a backing. In the experiments, four parameters—welding current, shielding gas flow rate, welding gap, and welding speed—were set as variables. Multiple experiments were conducted by varying the combinations of these parameters to observe how different settings influence the quality of the weld seam. Each parameter combination was tested three times to ensure data accuracy, with detailed data presented in Table 4.

**Table 4.** Experimental table of welding parameters.

Number	Main Welding Process Parameters				Number of Repetitions
	Welding Current	Shielding Gas Flow Rate	Welding Speed	Welding Gap	
01	205 A	0 L/min	47 mm/min	1.0 mm	3
02	175 A	15 L/min	47 mm/min	1.0 mm	3
03	255 A	15 L/min	30 mm/min	1.3 mm	3
04	170 A	15 L/min	47 mm/min	1.3 mm	3
05	205 A	15 L/min	47 mm/min	1.0 mm	3
06	205 A	15 L/min	47 mm/min	1.0 mm	3

The welding molten pool image is converted by video acquisition and post-processing, with a visual sampling frequency of 10 HZ, and the vibration signal and actual current value are acquired at a frequency of 1000 HZ and analyzed every 100 ms after welding. Adopting the same base material, welding consumables, equipment, and operating environment as that of the ship's hull plate, the simulated parts welding process test is carried out. The test adopted flat plate butt welding using 6 mm thickness Q235 steel, a type of mild steel. The angle of the open bevel is 60°, and the blunt edge measures 1.5 mm. The test simulated the actual situation of ship welding by varying the gap, welding speed, current, and gas flow. Obtained and collated to obtain 01 porosity, 02 incomplete penetration, 03 burnt through, 04 incompletely filled groove, 05 good, 06 weld misalignment a total of six types of quality types of welds of the combination of multi-source information data, the data information is shown in Table 5. Referring to the GB/T 3375-94 [21] welding terminology, to define the corresponding welding defects.

**Table 5.** Multi-source information data table.

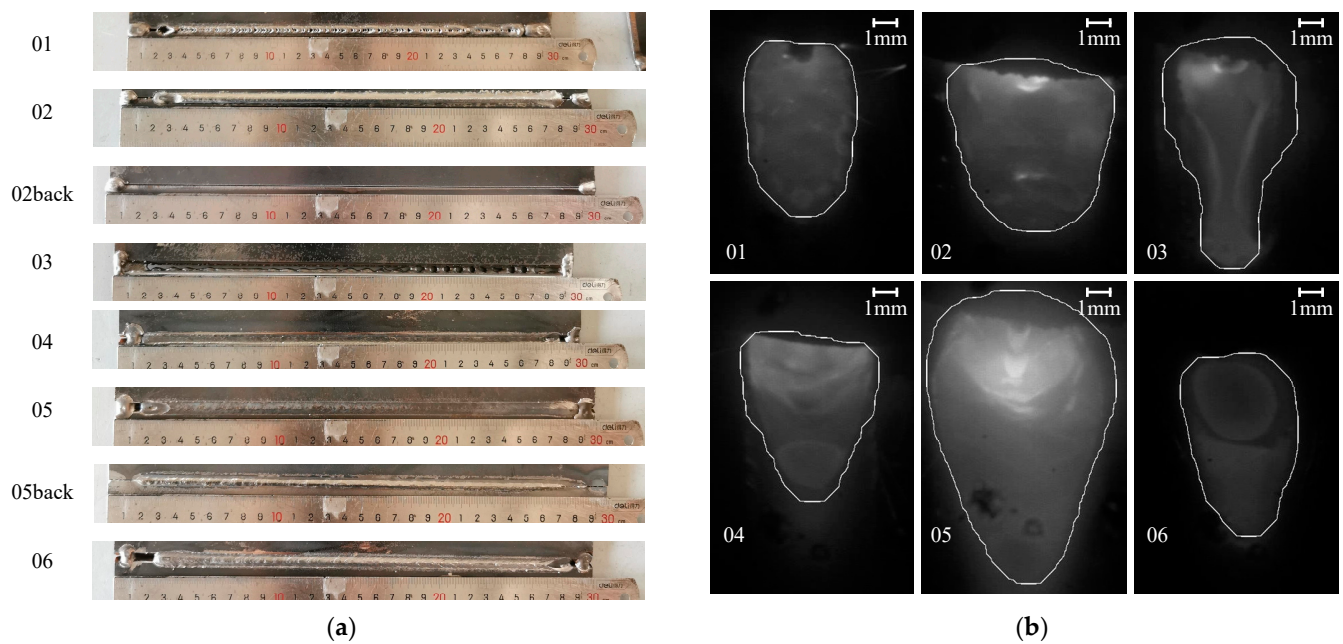
Number	Type	Data Sets n/Each
01	porosity	335
02	incomplete penetration	340
03	burn through	366
04	incompletely filled groove	359
05	favorable	358
06	weld misalignment	162

Porosity refers to the formation of cavities in the welding process where bubbles in the molten pool fail to escape during solidification and remain. Failure to weld through refers to the phenomenon in which the root of the joint is not completely penetrated during welding. It may also refer to situations where the depth of the weld does not meet the design requirements. Burn-through refers to a defect in which the molten metal flows out of the back of the bevel during the welding process, forming a perforation. Incomplete penetration refers to the formation of continuous or intermittent grooves on the surface of the weld due to insufficient filler metal. Weld deviation refers to the phenomenon that the center line of the weld gap and the center line of the weld toe is not centered. Overall, welding defects will weaken the structural strength of the ship, jeopardize safety, increase maintenance costs, and shorten its lifespan. Timely identification and repair of defects is of great significance to the safety and economic benefits of the ship.

#### 4.2. Melt Pool Image Analysis and Feature Extraction

The literature in [22] mentions a close relationship between weld quality and melt pool image features. Therefore, this paper presents melt pool images of porosity, incomplete penetration, burned-through, under-welded, good, and weld deviation. These images and their corresponding welds, as shown in Figure 5, were obtained by building a welding platform and designing experiments to simulate ship welding. From the physical weld images, it can be seen that each defect has a distinct appearance, 01 displays obvious

porosity; 02 appears similar to other defects from the front view, but the back view clearly shows no melt-through; 03 is clearly distinguishable from others, exhibiting signs of leakage and perforation; 04 shows a minor groove with the weld positioned below the workpiece when viewed from the back; 05 has the weld not below the workpiece from the front view, while the back shows some residual height and width; 06 shares similarities with other weld surfaces, but it can be differentiated as the weld is not centered within the gap. From the molten pool image, the shapes of these states are still somewhat differentiated, where the area of 05 is more round and full here, and the shape of 03 is gourd-like. By counting their appearance characteristics and shape features, the welding defect information system can be effectively enriched to provide a basis for automatic defect recognition. From the grey-level map of different states, the grey-level distribution varies across intervals. Thus, grey-level statistics can serve as a valuable addition to the defect information system.

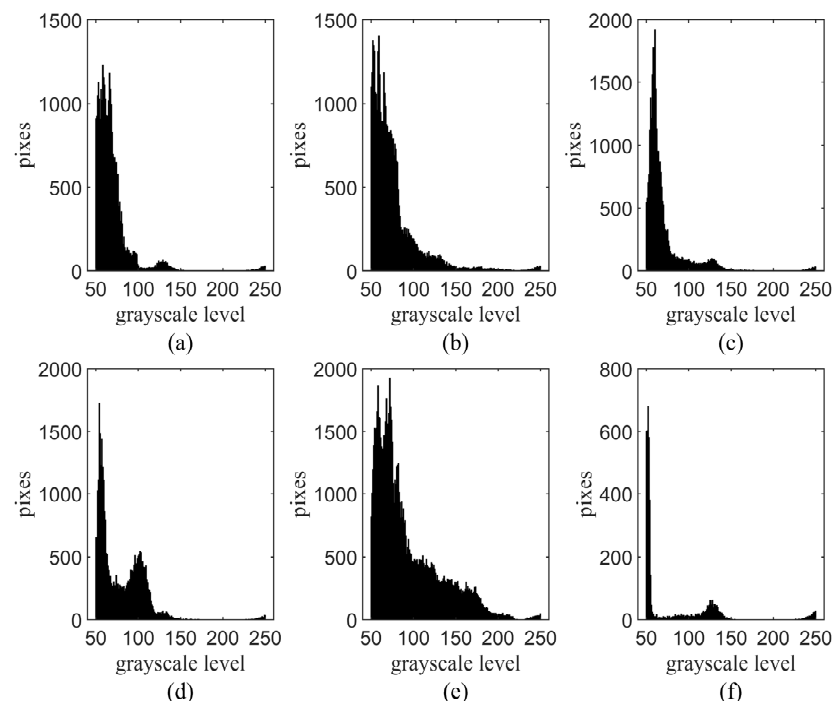


**Figure 5.** Welding defects and their corresponding melting pools. (a) physical drawing of the welded seam; (b) defect melt pool diagram.

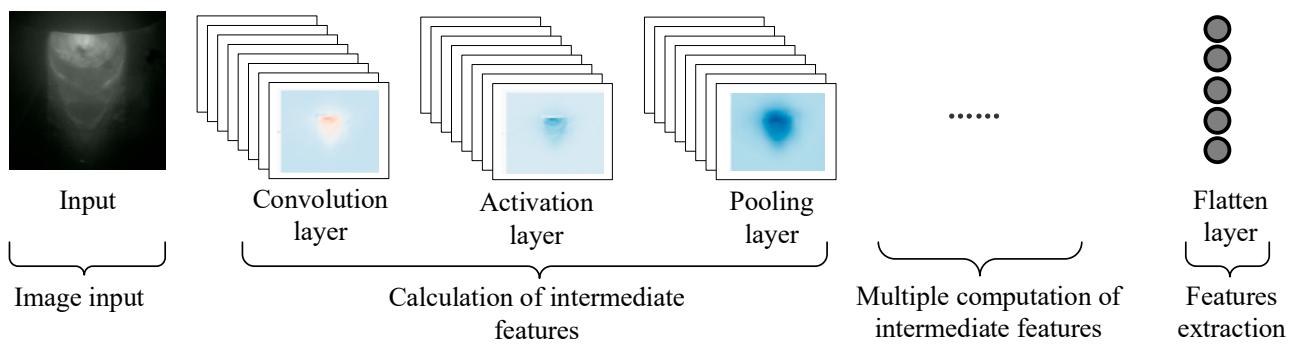
Combined with the analysis and generalization of the above-molten pool image characteristics and dimensional morphology, it is proposed to extract the continuous region with a grey value greater than 50 in the molten pool image as the profile of the molten pool and calculate the geometric dimensional information such as the maximum width of the molten pool,  $w$ , and the half-length of the molten pool,  $LT$ , under the six states of porosity, under-welded, burned through, under-welded, good, and weld deviation, respectively. The number of pixel points in the statistical interval, in order, is [50, 99] marked as P50, [100, 149] marked as P100, [150, 199] marked as P150, and [200, 255] marked as P200; the grey level map of each quality type is shown in Figure 6.

Convolutional neural networks (CNN) can effectively capture local and global features in an image through the combination of convolutional, pooling, and fully connected layers [23]. CNN can automatically extract key information such as edges, texture, and higher dimensional image information. In this paper, by constructing a melting pool image detection model of CNN and using the average pooling layer to extract the higher dimensional features of the melting pool image, we finally obtain the CNN of 128 molten pool attributes for combining with features from other information sources to construct an information decision system describing the defect recognition of molten pool images. The CNN processing principle and feature extraction process are shown in Figure 7. In this paper, a CNN for extracting molten pool images is constructed with specific parameters:

1. Input layer: accepts a  $32 \times 32 \times 3$  image as input, normalized using 'zerocenter'.
2. First convolution block:
  - Convolutional Layer 1: Uses  $32 \ 3 \times 3$  convolutional kernels in steps of  $[1, 1]$  with 'same' padding.
  - Batch Normalization 1: Normalizes 32 channels.
  - ReLU1 activation function.
  - Maximum Pooling Layer 1: Use a  $2 \times 2$  pooling kernel with steps of  $[2, 2]$ .
3. Second convolution block:
  - Convolutional Layer 2: Uses  $64 \ 3 \times 3$  convolutional kernels with 32 input channels, step size  $[1, 1]$ , and 'same' padding.
  - Batch Normalization 2: Normalizes 64 channels.
  - ReLU2 activation function.
  - Maximum Pooling Layer: Use a  $2 \times 2$  pooling kernel with steps of  $[2, 2]$ .
4. Third convolutional block:
  - Convolutional Layer 3: Uses  $128 \ 3 \times 3$  convolutional kernels with 64 input channels, step size  $[1, 1]$ , and 'same' padding.
  - Batch Normalization 3: Normalizes 128 channels.
  - ReLU3 activation function.
5. Feature convolution block:
  - Convolutional Layer 4: Uses  $128 \ 1 \times 1$  convolutional kernels with 128 input channels and a step size of  $[1, 1]$ .
  - Batch Normalization 4: Normalizes 128 channels.
  - ReLU4 activation function.
  - Global Average Pooling 4: Reducing the spatial size of features.
6. Spreading layer: extracting features



**Figure 6.** Grey scale chart for each weld quality. (a) welding quality01; (b) welding quality02; (c) welding quality03; (d) welding quality04; (e) welding quality05; (f) welding quality06.

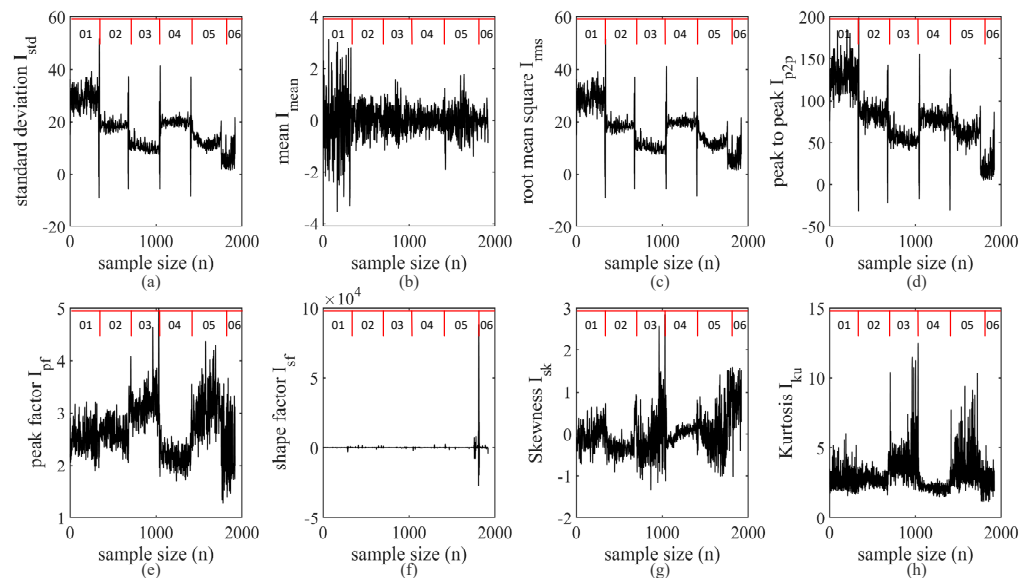


**Figure 7.** Schematic diagram of CNN feature extraction process.

### 4.3. Time-Domain Feature Analysis and Acquisition of Welding Process Parameters

The core parameters for quality control in the welding process are the welding process parameters, including welding current and arc voltage. However, in the actual production process, we often find discrepancies between the set values and the actual values. In this paper, we have collected the actual current values for different weld quality problems, such as porosity, incomplete penetration, burn-through, incompletely filled groove, satisfactory weld, and weld misalignment, respectively. Their actual current values are denoted as 01, 02, 03, 04, 05, and 06.

To delve deeper into these issues, we extracted eight data features from the actual current values, including standard deviation  $I_{std}$ , mean  $I_{mean}$ , root mean square  $I_{rms}$ , peak-to-peak  $I_{p2p}$ , peak factor  $I_{pf}$ , shape factor  $I_{sf}$ , skewness  $I_{sk}$ , and kurtosis  $I_{ku}$ . We then compared these characteristics with the actual current values of a normal (good) weld. As shown in Figures 8 and 9, we found that when there are problems such as porosity, under-welding, burn-through, under-welding, weld deviation, etc., their time-domain features are somewhat different compared to the eigenvalues of the ideal weld, which can be used as classifying features for defect identification.



**Figure 8.** Comparison of statistical characteristics of actual current for different weld qualities. (a) standard deviation; (b) mean; (c) root mean square; (d) peak to peak; (e) peak factor; (f) shape factor; (g) skewness; (h) kurtosis.

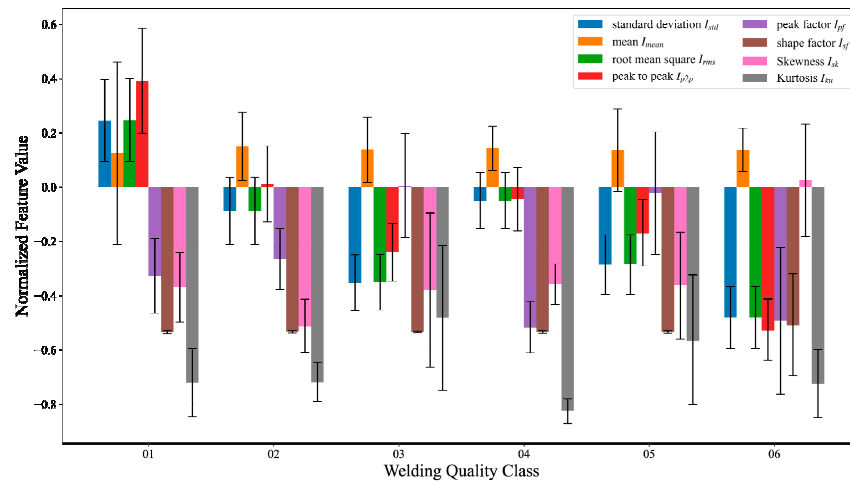


Figure 9. Current characterization error bar.

#### 4.4. Vibration Sensing Time Domain Feature Analysis and Acquisition

Vibration sensing is mostly used in troubleshooting machine tools or rotating machinery [24]; however, it is seldom used in welding. In fact, vibration is generated when welding, because we can hear the sound of the welding process, and the sound is generated from the vibration, through the monitoring of welding. The vibration signals during the process can be used to understand the fusion of the weld and the quality of the weld. Therefore, this paper attempts to use vibration sensors to detect the vibration signals emitted during the welding process, together with the current signals and molten pool images to achieve the identification of welding defects. Vibration signals for 01 porosity, 02 incomplete penetration, 03 burn-through, 04 incompletely filled grooves, 05 satisfactory weld, and 06 weld misalignment are collected. Due to environmental and equipment influences, the collected signals often contain noise. In this paper, we subject the signals to various processing methods, such as removing the DC component, wavelet thresholding and noise reduction (db3, 4-layer decomposition), signal frame analysis, and smoothing.

This helps in extracting the statistics of the vibration amplitude signals in the original X, Y, and Z directions. The statistical time-domain features of the signals with vibration amplitude in three directions are 16 dimensions, which are variance, maximum, peak, root-mean-square, absolute mean, peak indicator, skewness, waveform indicator, minimum, impulse indicator, margin indicator, square root magnitude, skewness indicator, mean values, and crag metrics. Since there are 16-dimensional features in one direction, combining the features from all 3 directions results in 48-dimensional features. For example, Figure 10 shows the time-domain waveform of the X-axis after removing the DC component and wavelet thresholding noise reduction.

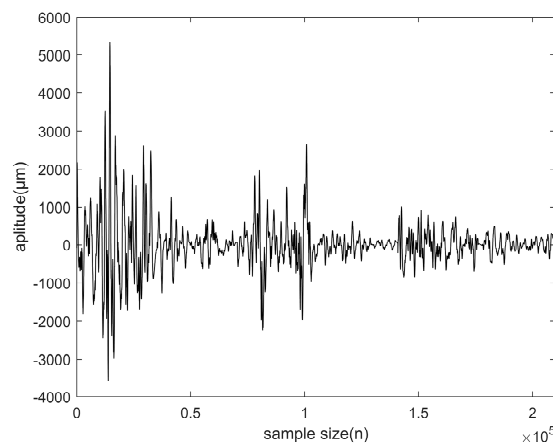


Figure 10. Vibration time domain signal.

#### 4.5. Welding Multi-Source Information Knowledge Modelling and Model Test Validation

To provide a more comprehensive and accurate description of the welding state, scholars have transitioned from relying on a single sensor to using multiple sensors. To maximize the use of data obtained from multiple sensors, they adopted multi-sensor information fusion techniques. Feature-level fusion is particularly effective due to its capability for significant information compression and its advantage in maximizing the information essential for decision analysis. This fusion first extracts features from each sensor data and then integrates these features to improve system performance, stability, and better understanding of the data. Therefore, in this paper, we propose to use feature level fusion to achieve the task of weld defect identification by obtaining 139 features of the molten pool image, 48 features of the vibration signal, and 8 features of the actual current value of the weld for six quality types of welds, namely, porosity, incomplete penetration, burnt-through, under-welded, good, and weld deviation. After the fusion process and utilizing the neighborhood rough set for feature fusion and attribute reduction, the resulting new decision system is used as an input for machine learning, and machine learning is used to achieve welding defect recognition. The multi-source information knowledge modeling process used in this paper is shown in Figure 11.

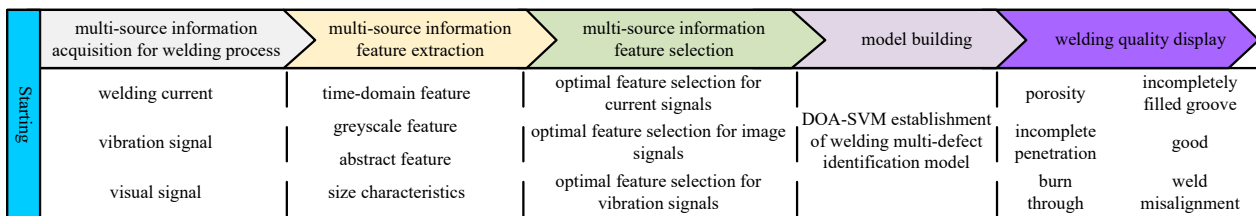


Figure 11. Multi-source information knowledge modeling process.

Using the same base material, welding consumables, equipment, and operating environment as the ship’s hull plate, we carried out welding process tests on simulated parts and obtained the collated test data to create a dataset. A total of 1920 sets of data were obtained. Among them, 335 groups are of porosity, 340 are of non-weld-through, 366 are of burn-through, 359 are of non-full weld, 358 are of good quality, and 162 are of weld deviation. Each of the three vibration signal directions (X, Y, and Z) has 16 features. Additionally, there are 8 current features and 139 visual image features, making a total of 195 features. Before diagnosing and identifying the defects, the experimental data are first preprocessed in the following three ways.

- To ensure that large-valued features do not overshadow the importance of small-valued features during the classification process, we normalized each feature using Equation (18).

$$x'_{ij} = \frac{x_{ij} - x_{j\min}}{x_{j\max} - x_{j\min}} \tag{18}$$

where  $x_{ij}$  represents the sample of the  $j$  column of the  $i$  row,  $x_{j\min}$  represents the minimum value of the  $j$  column, and  $x_{j\max}$  represents the maximum value of the  $j$  column.

- Outlier detection is then used to remove the outlier samples in the
- Due to the vast amount of data, direct training would be resource-intensive. Thus, attribute reduction is applied to this data table.

The dataset is divided into an 80% training set and a 20% test set. The feature selection algorithm presented in this paper is employed to generate feature subsets, and its effectiveness and the advantages of the DOA-SVM model are verified. After the parameter optimization of DOA on SVM, the highest classification accuracy is achieved at  $c = 1024$ ,  $g = 9.76 * 10^{-5}$ . The neighborhood radius  $\delta = 0.2$  and the kernel function RBF are set to train and test the support vector machine. Ten trials were conducted, and the average test

classification accuracy from these trials was used to evaluate each classification algorithm, including CART, SVM, and DOA-SVM. The trials were grouped into three sets:

- The sampled dataset undergoes no processing and is directly used for training and testing.
- The classical rough set theory’s attribute reduction algorithm is applied to approximate the attributes of the sampled dataset, after which the approximated set is used for training and testing.
- A neighborhood rough set-based attribute reduction algorithm is employed to approximate the attributes of the sampled dataset. Subsequently, support vector machines are trained and tested on this reduction set.

Table 6 demonstrates the effect of different feature selection methods and classifiers on classification accuracy and the number of features. It can be seen that NRS is most effective in reducing the number of attributes, reducing the features to only 12. Among all classifiers, DOA-SVM performs best with an average accuracy of 98.14%, confirming its superiority. When classical rough set (RS) was used for feature selection, the accuracy of all classifiers was reduced, probably because RS is only applicable to class attributes, which need to be discretized before processing continuous data, and some useful information may be lost. While using NRS, although the number of attributes is reduced significantly, the classification accuracy of DOA-SVM is still close to the level when using the original features, indicating that NRS can effectively retain the features useful for classification while reducing the number of features; in terms of the single training time consumption, the training of the simplified data in SVM reduces about 80% of the time consumption compared to the non-simplified features with the time of 0.11 s, which indicates that data dimensionality reduction is useful to improve diagnostic efficiency and reduce model complexity.

**Table 6.** Comparison of feature selection and classification results.

	CART	SVM	DOA-SVM	Feature No.	Single Training Time
Raw	95.81%	98.50%	99.22%	195	0.55 s
RS	94.53%	95.83%	96.22%	12	0.11 s
NRS	94.01%	98.43%	98.98%	12	0.11 s
Average	94.78%	97.59%	98.14%	75.67	0.26 s

Table 7 details the performance of the Neighborhood Rough Set (NRS)-based DOA-SVM model in identifying the six types of welding defects, namely, Porosity, Incomplete penetration, Burn-through, Incompletely filled groove, Good, and weld misalignment. Significantly, the model demonstrates an extremely accurate prediction for “burn-through” and “weld deviation”, with all relevant metrics reaching 1.0, while the model’s precision for “incompletely filled groove” is 0.989, which is slightly lower than 1.0, but with a recall of 1.0, which ensures the complete detection of such defective instances. For “Porosity” and “good”, although the recall is slightly lower than 1.0, which may result in a small number of samples being missed, all the samples recognized as being in these categories are correct, resulting in a precision of 1.0, while “incomplete penetration” is recognized with a precision of 0.937, which is the lowest of the six categories, but with a recall of 1.0 and an F1 score of 0.967, which is near 1.0. These performances show the high efficiency of the DOA-SVM model and the excellence of NRS in feature selection, which can meet the requirements of welding defect detection and identification.

**Table 7.** Statistics of NRS-based DOA-SVM experimental results.

Serial Number	Class	Accurate	Recall Rate	F1 Score
01	porosity	1.0	0.969	0.984
02	incomplete penetration	0.937	1.0	0.967
03	burn through	1.0	1.0	1.0
04	incompletely filled groove	0.989	1.0	0.994
05	good	1.0	0.957	0.978
06	weld misalignment	1.0	1.0	1.0

## 5. Conclusions and Outlook

### 5.1. Conclusions

1. A welding defect warning and identification system based on multi-source heterogeneous sensing data of molten pool images, current signals, and vibration signals is constructed by simulating the operation mode of experienced welders in the process of welding thin plates.
2. In the field of identification of welding defects, vibration sensors were used for the first time, and the importance of vibration signals in the identification process was confirmed in experiments.
3. A support vector machine classification method based on neighborhood rough sets is introduced to reduce the size and complexity of the problem and increase the speed of diagnosis. The feature dimensions are reduced from 195 to 12, significantly lowering the complexity of the problem. The training time for a single model is reduced from 0.55 s to 0.11 s, greatly cutting resource consumption and improving diagnostic speed. Experimental results show that the method has high generalization potential.
4. This paper presents a new method that combines NRS with DOA-SVM to quickly and accurately distinguish between five types of defects and good welds in arc welding. The identification accuracy of this method is at least 98%, an improvement of approximately 4.97% compared to CART and 0.55% compared to standard SVM. These results confirm the research and application value of the method.

### 5.2. Outlook

Currently, our research is in the experimental phase and has not yet been applied on an industrial scale. Future work will focus on developing corresponding software and hardware systems to facilitate industrial applications. This includes:

1. Developing software interfaces suitable for different industrial welding scenarios to integrate multi-sensor data.
2. Optimizing hardware systems to adapt to various operating conditions in industrial environments.
3. Collaborating with industry partners to conduct large-scale field tests to verify the system's performance and reliability in actual production.

Through these efforts, we aim to translate the research outcomes into practical industrial applications, thereby improving welding quality and reducing production costs.

**Author Contributions:** Conceptualization, Q.L.; Data curation, X.H. and B.L.; Formal analysis, X.X., X.H. and Z.P.; Funding acquisition, Z.F.; Investigation, X.Z., X.X. and X.L.; Methodology, X.Z.; Project administration, Z.F. and B.L.; Resources, X.L.; Supervision, Z.F., Z.P. and Q.L.; Validation, X.X. and X.L.; Visualization, X.Z. and X.H.; Writing—original draft, X.Z.; Writing—review & editing, Z.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China under Grant No. 52261044 and the Guangxi Science and Technology Major Project under Grant No. Guike AA23062037.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data can be obtained by contacting the author Xianping Zeng (18523951688@163.com). The data are not publicly available due to privacy concerns.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Yuyue, L. Research on process optimization and parameter control of large-scale mechanical welding. *Mach. China* **2023**, *2023*, 51–54.
2. Xin, L.; Lili, Z. Research progress of intelligent robot welding technology. *Scientist* **2017**, *5*, 19–20+48.

3. Kun, Z.; Zongxuan, Z.; Ye, L.; Zhengjun, L. Multi-sensor data collaborative sensing algorithm for aluminum alloy TIG welding pool state. *Trans. China Weld. Inst.* **2022**, *43*, 50–55+116. [[CrossRef](#)]
4. Wang, J.; Wen, X.; He, Y.; Lan, Y.; Zhang, C. Logging curve prediction based on a CNN-GRU neural network. *Geophys. Prospect. Pet.* **2022**, *61*, 276–285. [[CrossRef](#)]
5. Chandra, M.A.; Bedi, S. Survey on SVM and their application in image classification. *Int. J. Inf. Technol.* **2021**, *13*, 1–11. [[CrossRef](#)]
6. Ma, F.; Li, X. Landslide displacement prediction model using improved SSA-KELM coupling algorithm. *Sci. Technol. Eng.* **2022**, *22*, 1786–7963.
7. Zhang, X.; Liu, C.; Xue, L.; Zeng, H. Simultaneous feature selection and SVM parameter by using artificial bee colony algorithm. In Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering, Xiamen, China, 21–23 October 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 1737–1745. [[CrossRef](#)]
8. Alikhan, J.S.; Alageswaran, R.; Amali, S.M.J. Dingo optimization based network bandwidth selection to reduce processing time during data upload and access from cloud by user. *Telecommun. Syst.* **2023**, *83*, 198–208. [[CrossRef](#)]
9. Prasuna, P.; Ramadevi, Y.; Babu, A. A two level approach to discretize cosmetic data using Rough Set Theory. *Int. J. Comput. Technol.* **2015**, *14*, 6147–6152. [[CrossRef](#)]
10. Li, W.; Xia, S.; Chen, Z. A Fast Attribute Reduction Algorithm of Neighborhood Rough Set. In Proceedings of the 2021 13th International Conference on Knowledge and Smart Technology (KST), Bangsaen, Thailand, 21–24 January 2021; pp. 43–48. [[CrossRef](#)]
11. Hu, Q.; Yu, D.; Xie, Z. Numerical Attribute Reduction Based on Neighborhood Granulation and Rough Approximation. *J. Softw.* **2008**, *19*, 640–649. [[CrossRef](#)]
12. Gao, X.; Li, G. A KNN model based on manhattan distance to identify the SNARE proteins. *IEEE Access* **2020**, *8*, 112922–112931. [[CrossRef](#)]
13. Chervonenkis, A.Y. Early history of support vector machines. In *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 13–20. [[CrossRef](#)]
14. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
15. Burges, C.J. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [[CrossRef](#)]
16. Liu, C.; Han, J. Application of support vector machine based on neighborhood rough set to sewage treatment fault diagnoses. *J. Gansu Agric. Univ.* **2013**, *48*, 176–180. [[CrossRef](#)]
17. Liu, F.T.; Ting, K.M.; Zhou, Z.-H. Isolation forest. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 15–19 December 2008; pp. 413–422. [[CrossRef](#)]
18. Peraza-Vázquez, H.; Peña-Delgado, A.F.; Echavarría-Castillo, G.; Morales-Cepeda, A.B.; Velasco-Álvarez, J.; Ruiz-Perez, F. A bio-inspired method for engineering design optimization inspired by dingoes hunting strategies. *Math. Probl. Eng.* **2021**, *2021*, 9107547. [[CrossRef](#)]
19. Almazán-Covarrubias, J.H.; Peraza-Vázquez, H.; Peña-Delgado, A.F.; García-Vite, P.M. An improved Dingo optimization algorithm applied to SHE-PWM modulation strategy. *Appl. Sci.* **2022**, *12*, 992. [[CrossRef](#)]
20. Milenković, B.; Jovanović, Đ.; Krstić, M. An application of Dingo Optimization Algorithm (DOA) for solving continuous engineering problems. *FME Trans.* **2022**, *50*, 331–338. [[CrossRef](#)]
21. GB/T 3375-1994; Welding Terminology. China Standards Press: Beijing, China, 1994.
22. Wang, K.; Shen, Y.; You, F.Q.Q. MAG weld pool image features and available information analysis. *Trans. China Weld. Inst.* **2006**, *11*, 53–56+115. [[CrossRef](#)]
23. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *60*, 84–90. [[CrossRef](#)]
24. Jose, M.; Kumar, S.S.; Sharma, A. Vibration assisted welding processes and their influence on quality of welds. *Sci. Technol. Weld. Join.* **2016**, *21*, 243–258. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.