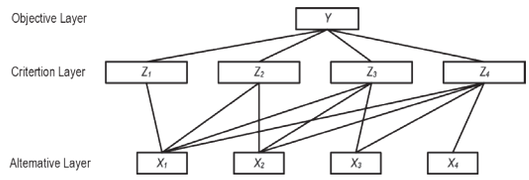
Ultrasonic metal welding quality prediction model based on MHSA–LSTM–AHP



Modelo de predicción de la calidad de la soldadura ultrasónica de metales basado en MHSA-LSTM-AHP

Xihao Fan[1,2], Hua Zhang[1,2,*] , Mei Zou[1,2], Ruijuan Wang[1,2] and Jingchi Zhang[3]

[1] Hubei University of Arts and Science. Hubei Key Laboratory of Power System Design and Test for Electrical Vehicle – 441053 Xiangyang, Hubei (China).
[2] Hubei University of Arts and Science. School of Automotive and Traffic Engineering – 441053 Xiangyang, Hubei (China).
[3] University of New South Wales. School of Automotive and Traffic Engineering - NSW 2033 Sydney (Australia).

DOI: https://doi.org/10.52152/D11433 | Received: 11/mar/2025 • Reviewing: 13/mar/2025 • Accepted: 08/jun/2025

To cite this article: FAN, Xihao; ZHANG, Hua; ZOU, Mei; WANG, Ruijuan; ZHANG, Jingchi. ULTRASONIC METAL WELDING QUALITY PREDICTION MODEL BASED ON MHSA–LSTM–AHP. *DYNA* July August 2025. Vol. 100, n.4, pp. 363-370 | https://doi.org/10.52152/D11433

FUNDING

This work was supported by Open Fund of Hubei Key Laboratory of Electric Vehicle Power System Design and Testing (Grant No. ZDSYS202517).

RESUMEN

- La tecnología de soldadura ultrasónica de metales (UMW) se utiliza ampliamente en industrias como la fabricación de vehículos eléctricos debido a su eficiencia, bajo aporte de calor y capacidad para unir metales diferentes. Sin embargo, la calidad de la soldadura de la tecnología UMW es susceptible a diversos parámetros del proceso, por lo que garantizar la consistencia es un reto. En este estudio se propuso un modelo de predicción de la calidad que combina el proceso jerárquico analítico (AHP), la memoria a corto y largo plazo (LSTM) y la autoatención múltiple (MHSA) para evaluar con precisión la calidad de las uniones de la tecnología UMW. El proceso UMW se analizó en cuatro etapas basadas en las operaciones de producción actuales, y se extrajo información característica a partir de los datos característicos del proceso muestreados por el equipo de inspección en línea. Se desarrolló un marco jerárquico basándose en los principios del AHP para dilucidar las interrelaciones entre los datos del proceso en tiempo real, las características del proceso y la calidad conjunta de la UMW. En este estudio se estableció un modelo de predicción de la calidad MHSA-LSTM-AHP aprovechando las ventajas de LSTM y MHSA en el aprendizaje de dependencias temporales. Se realizó un análisis comparativo utilizando la red neuronal de retropropagación optimizada por algoritmos genéticos (GA-BP) y los modelos LSTM-AHP para evaluar el rendimiento del modelo propuesto. Los resultados indican que el modelo propuesto funciona bien en la predicción de la resistencia a la tracción y la resistencia de contacto, con un error medio del 3,21 % y el 3,7 %, respectivamente. Este estudio puede proporcionar una referencia satisfactoria para la construcción de un sistema de control de calidad en línea y la optimización del proceso UMW.
- **Palabras clave:** Soldadura ultrasónica de metales, modelo de predicción, memoria a corto y largo plazo, proceso jerárquico analítico, autoatención múltiple.

ABSTRACT

Ultrasonic Metal Welding (UMW) technology is widely used in industries such as electric vehicle manufacturing owing to its efficiency, low heat input, and ability to join dissimilar metals. However, the weld quality of UMW technology is susceptible to various process parameters; thus ensuring consistency is challenging. A quality prediction model that combines the Analytic Hierarchy Process (AHP), Long Short-term Memory (LSTM), and Multi-head Self-attention (MHSA) was proposed in this study to evaluate the joint quality of UMW technology accurately. The UMW process was analyzed in four stages based on current production operations, with characteristic information extracted from the process characteristic data sampled by the in-line inspection equipment. A hierarchical framework was developed by drawing on the principles of the AHP to elucidate the interrelationships between the real-time process data, the process characteristics, and the UMW joint quality. An MHSA-LSTM-AHP quality prediction model was established in this study by harnessing the advantages of LSTM and MHSA in learning temporal dependencies. A comparative analysis was conducted by using the Genetic Algorithm-optimized Backpropagation (GA-BP) Neural Network and LSTM-AHP models to evaluate the performance of the proposed model. Results indicate that the proposed model performs well in predicting tensile strength and contact resistance, with a mean error of 3.21% and 3.7%, respectively. This study can provide a satisfactory reference for the construction of an online quality monitoring system and the optimization of the UMW process.

Keywords: Ultrasonic Metal Welding, Prediction Model, Long Short-term Memory, Analytic Hierarchy Process, Multi-head Self-attention.

1. INTRODUCTION

Ultrasonic Metal Welding (UMW) is a solid-state welding technology that uses high-frequency ultrasonic energy. Owing to its low dependency on the intrinsic properties of materials and ability to minimize energy loss during the formation of the contact zone in welded compounds, UMW is widely used in the production of automotive components, such as lithium-ion battery tabs, wire terminals, and electronic devices[1]-[3]. Despite such advantages, UMW technology faces a major challenge: lack of reliable nondestructive technologies for evaluating the weld quality [4-5]. The current industry standards for assessing the weld quality refer mainly to the "Performance Specification for Ultrasonically Welded Wire Terminations," which was published by USCAR (Unit-

ed States Council for Automotive Research) in 2016 [6]. In the specification, tensile strength testing is conducted on weld joints to evaluate their mechanical properties, and contact resistance testing is performed to assess their electrical properties. Nevertheless, quality assessment methods are destructive and not suitable for actual manufacturing processes [7-8]. Thus, in research communities, growing emphasis is being placed on the development of reliable mathematical models to predict the quality of UMW joints, coupled with the implementation of nondestructive evaluation techniques.

Although the existing studies established foundational frameworks for quality prediction [9], the practical utility of such frameworks remains constrained by their limited accuracy under dynamic manufacturing conditions. Advancements in sensor technology have enabled the acquisition of high-resolution process data, which exhibit complex nonlinear relationships with weld quality outcomes. The incorporation of such parameters as input makes quality prediction models highly effective but complex. Thus, critical challenges lie in the exploration of the relationships between sensor-derived features and final quality indicators and the development of a robust model to accurately capture the interdependencies among the parameters in the UMW process to improve joint quality prediction.

This study extracts critical features from sensor signals and proposes a novel relationship model that links UMW parameters to joint quality. The model integrates the Analytic Hierarchy Process (AHP), Multi-head Self-attention (MHSA), and Long Short-term Memory Network (LSTM). This study examines the performance of the proposed model by comparing it with existing models. The objective of this study is to illustrate the interdependencies among the parameters during the UMW process clearly and predict the weld quality accurately. By achieving the objectives, this study can provide a theoretical foundation and practical reference for the optimization of UMW quality and the precise control of the welding system.

2. STATE OF THE ART

Current methods for UMW quality assessment can be categorized into two approaches: mechanism analysis and data-driven approaches. Mechanism analysis typically relies on finite element modeling software to simulate the welding process and investigate the impact of process parameters on the weld quality [10]. Shen et al. [13] developed a three-dimensional finite element model to simulate the response of composite materials under UMW conditions. The authors validated the model by using experimental data, including welding pressure, temperature, and weld geometry, which were measured during the actual welding process. Li Huan [14] used the ANSYS finite element software to examine the effect of the welding amplitude on UMW by obtaining the plastic strain and temperature changes of the weld under different amplitude conditions. Simulation models can provide a satisfactory explanation of the physical welding process, but the establishment and validation of finite element models can become challenging when the weld conditions change.

Meanwhile, data-driven approaches leverage large-scale welding experimental datasets to perform feature analysis and establish predictive models that link process parameters and measured factors to weld quality through advanced techniques, such as data mining and machine learning. Yao et al. [15] used the response surface methodology to examine the optimal UMW process parameters for copper wire and aluminum sheets. The authors utilized a Box–Behnken design to perform the analysis and constructed regression models to optimize the welding parameters. Sarraf [16] used an artificial neural network to predict the strength of Cu-Al joints and employed pressure, the welding duration, and the vibration amplitude as the input parameters. Mongan et al. [17] integrated genetic algorithms and neural networks to build a model for predicting the shear strength of UMW aluminum alloys and used the welding energy, amplitude, and clamping force as the input parameters and the peak power as the process feedback for closed-loop prediction. Although the methods demonstrated predictive capabilities for weld quality under controlled conditions, they exhibit a critical limitation in practical UMW application. The aforementioned studies focused predominantly on modeling the relationships between process parameters and weld quality but neglected the influence of dynamic and uncontrollable process state parameters that can significantly affect joint consistency in industrial environments.

With the advancement of sensor technology, high-precision sensors have been developed and applied to obtain welding process characteristic information. Featured information has been derived from acquired sensing data, and data-driven approaches have been used in weld quality prediction. For instance, Lee et al. [18] identified the features of welding power and horn displacement to predict joint quality, and Choi [19] extracted nine feature variables from energy and weld sonotrode displacement signals, including the peak, mean, and standard deviation, during the welding process. The author used the features to classify joint quality into three categories: sufficient welding, incomplete welding, and overwelding. Moreover, a support vector machine model was developed to classify weld quality, which achieved an impressive accuracy of 98%. Guo et al. [20] performed complex feature selection and used the Statistical Process Control-M-distance algorithm to classify poor welds in battery tab welding. Hong [21] analyzed power and Linear Variable Differential Transformer signals and used convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to classify weld quality. The aforementioned methods can predict joint quality types effectively but cannot determine quantitative quality indicators accurately, such as tensile strength and contact resistance.

Samir [22] utilized response surface methodology to establish the relationships between three UMW process parameters (welding pressure, time, and amplitude) and three measured factors (power, force, and energy) in aluminum alloy UMW. However, response surface methodology struggles to represent the highly nonlinear relationship between the process parameters and the measured factors accurately. Müller et al. [23] used sensors to obtain three signals during welding: mechanical vibration, the sonotrode penetration depth, and the internal welding power. The authors processed and converted the collected data into a random forest model to estimate the shear strength of the weld joint. In another study [24], the authors considered the difficulty of measuring external sensor data in industrial environments and measured only high-frequency power signals in a welding system. Various machine learning models have been used for quality prediction, with the best-performing model achieving a root mean square error (RMSE) of 89 N. However, the models do not consider the varying effects of process parameters at different time sequences on the weld quality. Schwarz [25] analyzed signal curves from a welding machine, a sonotrode, and a welding table during the welding process and extracted over 700 features. The author conducted linear regression and multilayer perceptron regression to predict the tensile shear strength of the sonotrode, but the ex-

tensive number of extracted features introduced overfitting risks in the prediction model. Chen [26] applied variational mode decomposition and particle swarm optimization to extract sensor vibration signals. Given the high-frequency nature of vibration signals and the low-frequency characteristics of power and pressure signals, the author developed a spectrum-customized denoising diffusion probabilistic model to generate time-series data. Subsequently, the author employed a 17-layer MobileNetV2 network to establish a UMW quality prediction model. However, the author did not consider the strong coupling between the vibration signals and the power signals, which complicated the model's computational process. Wu et al. [27] proposed a novel end-to-end online UMW quality prediction method based on sensor fusion and deep learning algorithms. The authors used discrete wavelet transform to convert UMW process signals into two-dimensional images and obtained 200 experimental samples. Then, the authors used the samples to train a ResNet20 model, which achieved an R^2 value of 0.43. Nevertheless, the extracted features did not undergo stage-specific analysis, which rendered them inadequate to represent the complete welding process precisely.

The aforementioned achievements encompass studies on the correlation between UMW quality and process parameters, as well as interrelationships among process parameters. However, several issues remain unresolved. UMW is a multistage sequential process, and the parameters in each stage can influence the weld quality. Research on the relationship between process parameters and the weld quality in each individual stage is lacking. In addition, the features selected for weld quality prediction models are highly complex, and coupling relationships between sensor-derived features and process parameters have yet to be considered adequately. The gaps highlight the need for focused and nuanced approaches to fully understand and optimize the UMW process.

To address the above issues, first, this study analyzes the process parameters in each UMW stage, as well as the actual production situation. Second, this study uses external sensors to extract the characteristic information, such as temperature and the displacement of the sonotrode, in each stage. Third, this study employs the hierarchical concept of the AHP to represent the nonlinear relationships between the process parameters, the process characteristic data, and the weld quality indicators. Fourth, this study introduces the LSTM neural network to solve the problem of dependence of long-distance time-series data. Fifth, this study uses MHSA to evaluate the weights of different time-series positions and enable the model to focus on the process parameter factors in each stage that can impact the weld quality and improve the weld quality prediction accuracy. Last, this study develops an MHSA-LSTM-AHP model that can clearly express the time-series relationships between the process parameters, the process characteristic data, and the quality indicators to accurately predict the weld quality.

The remainder of this study is organized as follows: Section 3 analyzes the UMW process and introduces the proposed UMW weld quality prediction model, Section 4 presents the training of the proposed predictive model by using welding experiment samples and compares the model with other predictive models to evaluate its efficiency and performance, and Section 5 summarizes the findings and relevant conclusions.

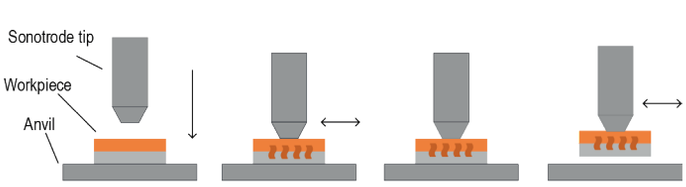


Fig. 1. Stages of UMW Process.

3. METHODOLOGY

3.1. UMW PROCESS

A real UMW process can be segmented to four stages as shown in Fig. 1. The features of the sensor signals used in each stage are analyzed to comprehensively investigate the influence of the process parameters in each UMW stage on the final weld quality.

Initial Contact Stage

During the initial contact stage (Fig. 1a), under the action of a cylinder, the sonotrode of the welding equipment applies a certain amount of pressure to the workpiece placed underneath to bring the upper and lower metal workpieces in close contact. Pressure is maintained until the pressure-holding phase ends. In this period, the key process parameter is the clamping pressure magnitude, *P*.

Ultrasonic Vibration Stage

During the ultrasonic vibration stage (Fig. 1b), the ultrasonic generator generates high-frequency electrical signals, which are converted into mechanical vibrational energy by a transducer. Then, the booster amplifies the vibration amplitude and transfers the energy to a sonotrode. The main parameters that can affect the weld quality in this stage are the set amplitude *C*1, power *W*1, and welding time *T*1.

Pressure-holding Stage

During the pressure-holding stage (Fig. 1c), the ultrasonic vibration ceases. Nevertheless, the sonotrode maintains constant pressure on the workpiece. Under the continuous pressure, the welded joint gradually cools and solidifies, giving rise to a stable solid-state connection. Throughout this stage, the clamping pressure remains constant, and the pressure-holding time *T*2 exerts an influence on the weld quality.

Post-welding Stage

During the post-welding stage (Fig. 1d), slight vibration is typically applied to the sonotrode after the pressure-holding stage to prevent the workpiece from sticking to the sonotrode. This vibration can break the possible micro connections between the workpiece and the sonotrode and thus enable their separation. This stage is similar to the ultrasonic vibration stage, and the main parameters are the set amplitude *C*2, power *W*2, and post-welding treating time *T*3.

The process parameters in each stage can be expressed as follows:

$$X=(X_1,X_2,X_3,X_4)^T$$
$$\text{Where } X_1=P,\ X_2=(C1,W1,T1)^T,\ X_3=T_2, \text{ and } X_4=(C2,W2,T3)^T. \quad (1)$$

In UWM production, engineers will set appropriate process parameters to suit specific materials and dimensions. This includes setting the center frequency of the transducer on the welding machine and the dimensions of the welding head, among other relevant parameters. Throughout the production process, the real-time process data is collected using real-time Data Acquisition Devices (DAQ), which log the actual production process data.

To ensure the rationality of the production process parameters, the in-line inspection equipment such as the infrared thermography and the laser displacement tester are generally employed to measure

the process characteristic data such as the amplitude of the welding head, the downward displacement, and the workpiece temperature during the welding process. To collect the process characteristic data, the laser displacement sensor KEYENCE LG-K5001 was used to measure the amplitude of the welding head and its downward displacement, and the infrared thermometer GAODE IPT640M was used to measure the temperature of the weldments. The general variation trends of the three types of process characteristic data are shown in Fig. 2 (see section: supplementary material).

The displacement depth exhibits a trend of rapid increase, slow increase, and rapid decrease. The rapid growth is caused by the rapid downward movement of the sonotrode under the pneumatic device, and the peak depth value *D*1 in this stage is selected as a feature variable. Subsequently, the ultrasonic vibration energy induces microscopic deformation in the metal, which leads to a slow increase in the sonotrode depth. For this stage, the peak value, the average value, and the standard deviation (*D*2, *D*3, *D*4) of the sonotrode depth are selected. During the pressure-holding phase, the sonotrode depth continues to rise slightly under the applied pressure, and the peak depth value *D*5 in this stage is selected.

During the initial contact stage, in which the workpiece temperature remains at ambient levels, the mean temperature *E*1 is identified as a key feature variable. Upon the initiation of ultrasonic vibration, the rapid temperature rise induced by frictional heating is characterized by the peak temperature, the mean temperature, and the standard deviation (*E*2, *E*3, *E*4), which are selected as the critical feature variables for the ultrasonic vibration stage. The temperature gradually declines after the cessation of the ultrasonic vibration. The mean temperatures (*E*5, *E*6) during the pressure-holding stage and the post-welding stage are defined as the representative feature variables during the stages.

The amplitude signal of the sonotrode emerges only in the ultrasonic vibration stage and the post-processing stage, which exhibits a pattern of sharp increase, stabilization, and then sharp decrease. A strong correlation exists between the amplitude value information and weld quality. Therefore, the peak value named *S*1, the mean value *S*2, and the standard deviation *S*3 of the sonotrode amplitude in the ultrasonic vibration stage are selected as the feature variables. Similarly, for the post-processing stage, the peak value *S*4, the mean value *S*5, and the standard deviation *S*6 of the sonotrode amplitude are selected.

In the subsequent context, the eigenvalues extracted from the aforementioned process characteristic data are denoted as *Z*, as follows:

$$Z=(Z_1,Z_2,Z_3,Z_4)^T$$
$$\text{Where } Z_1=(D1,E1)^T,\ Z_2=(D2,D3,D4,E2,E3,E4,S1,S2,S3)^T,\quad (2)$$
$$Z_3=(D5,E5)^T,\text{ and } Z_4=(E6,S4,S5,S6)^T.$$

After the welding process, the offline inspection equipment is used to inspect the quality indicators of the welded joints. According to [6], the joint tensile shear strength *F* and the contact resistance *R* are used to evaluate the weld quality, which are represented by *Y*.

$$Y=(F,R)^T \quad (3)$$

3.2. PRINCIPLES OF QUALITY PREDICTION

3.2.1. Quality assessment based on AHP

During the UMW process, process parameters can determine the changes in the process characteristic data, and there is a high degree of coupling between them. As is known, the process characteristic data is closely related to welding quality. However, in industrial production, due to cost considerations, existing welding systems typically do not measure the process characteristic data, which is crucial for subsequent optimization and precise control. For instance, when adjusting process parameters, it is necessary to ensure that process information parameters are within an appropriate range based on the process characteristics. Therefore, when establishing a predictive model, in addition to considering the relationship between process parameters and welding quality, the relationship between the process data and the process characteristic data must also be taken into account.

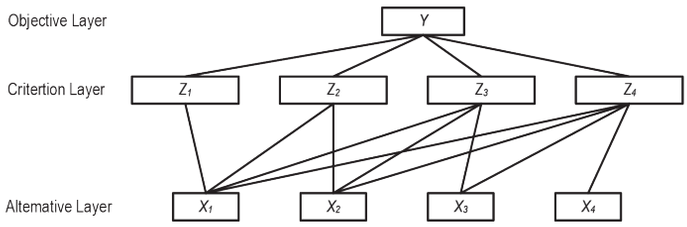


Fig.3. Hierarchical Structure of the process parameters, the process characteristic data, and the weld quality.

The hierarchical structure of the AHP [28-29] can enable the decomposition of the UMW quality assessment into three layers: objective, criterion, and alternative. The decomposition can clarify and simplify the dependency and coupling relationships between the factors at each level. The current process parameters will affect the process characteristic data in subsequent stages to varying degrees without affecting those in the preceding stages. Therefore, the *X*, *Z*, and *Y* extracted in Section 3.1 can be represented by the relationship shown in Fig. 3.

3.2.2. LSTM

Although the model in Fig. 3 clearly shows the relationships between the variables in the three layers during the welding process, the traditional AHP struggles to handle high-dimensional data, such as *X* and *Z*. Thus, other advanced computational methods that can handle high-dimensional data and learn temporal dependencies should be introduced to improve the structure. Commonly used methods include CNNs [30], RNNs, and LSTM [31]. Among them, LSTM is an RNN improvement and uses memory units to retain long-term information across a time series and discard useless information to effectively learn the temporal relationships in the high-dimensional data. The basic structure of an LSTM memory cell is shown in Fig. 4 (see section: supplementary material).

The memory cell consists mainly of the cell state C_t, the input gate i_t, the output gate O_t, and the hidden state h_t, which interact with one another through sigmoid and tanh functions. The next section provides a detailed introduction to each component.

The forget gate is used to forget information selectively, which outputs a value between 0 and 1 through the sigmoid function, as follows:

$$f_t=\sigma(W_f\cdot[h_{t-1},x_t]+b_f) \quad (4)$$

Where x_t is the input vector, h_{t-1} denotes the hidden vector of the previous layer. W_f and b_f denote the weights and bias values of the inputs, respectively.

The input gate is composed of two parts. The sigmoid function determines the values to be updated, whereas the tanh function creates a new candidate vector of values. The equations are as follows:

$$i_t=\sigma(W_i\cdot[h_{t-1},x_t]+b_i) \quad (5)$$

$$\bar{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (6)$$

In which i_t represents whether the value needs to be updated or not, and $\bar{C}$ indicates a vector of new candidate values that will be added into the LSTM memory.

The output gate is responsible for determining the part of the memory cell state to be outputted to the hidden state. The calculation equations are as follows:

$$O_t = \sigma(W_O \cdot [h_{t-1}, x_t] + b_O) \quad (7)$$

$$h_t = O_t * \tanh(C_t) \quad (8)$$

Where O_t is the output value, and h_t is its representation as a value between −1 and 1.

The update of the cell state C_t is derived from the previous cell state C_{t-1} based on forgetting information f_t and the information to be added i_t.

$$C_t = f_t * C_{t-1} + i_t * \bar{C}_t \quad (9)$$

Where f_t is the results of the forget gate, which is a value between 0 and 1 where 0 indicates completely get rid of the value; whereas, 1 implies completely preserve the value.

3.2.3. MHSA

As mentioned above, the self-attention mechanism can assess the contribution rate of different sample historical inputs to the output, according to the input data, and assign additional weights to the key feature information. When incorporated into a model, MHSA, which is a self-attention mechanism improvement, can enhance the prediction ability of LSTM [32] and determine the importance of each process parameter to the final weld quality and thus provide valuable insights for future process optimization. The calculation steps are described below.

As shown in Fig. 5 (see section: supplementary material), input X is mapped through three learnable parameter matrices, namely, W_h^q, W_h^k and W_h^v, to obtain the query vector Q_h, the key vector K_h, and the value vector V_h, where the subscript h represents the attention head index. Then, the dot product of Q and K is used to measure the similarity between the features. Subsequently, the softmax function is applied to the attention scores to calculate the attention weight, as follows:

$$a_i = \text{softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right) \quad (10)$$

Where Q_i and K_i are matrices of queries and keys, d_k is the vector dimension, which is used to control the magnitude of the dot-product result. Weight a_i is applied to vector v_i to obtain the output of each head,

$$head_i = a_i V_i \quad (11)$$

The outputs of all the heads are concatenated and transformed through the learnable output weight parameter matrix W^O to obtain the final feature representation.

$$MultiHead = Concat(head_i, \ldots, head_h) W^O \quad (12)$$

3.3. MHSA-LSTM-AHP MODEL

By leveraging the strengths of the AHP in managing complex relationships, along with the capabilities of MHSA and LSTM for time-series prediction, this study introduces the MHSA-LSTM-AHP model for weld quality prediction. The model structure is shown in Fig. 6.

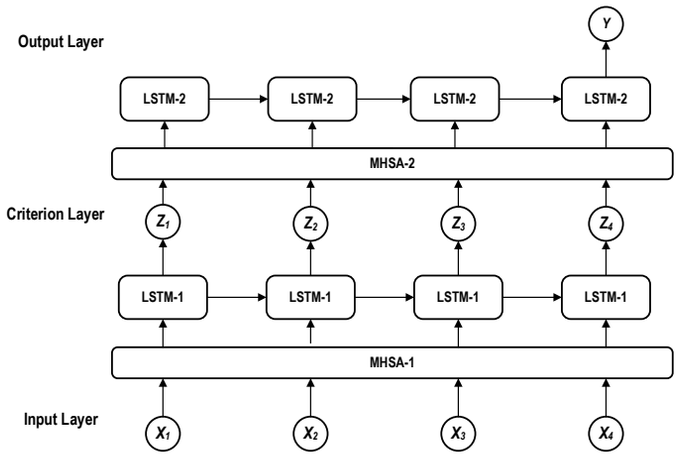


Fig. 6. Hierarchical Structure of MHSA-LSTM-AHP Model.

The model consists of seven layers, which are listed in Table 1. A description of each component of the proposed model is provided below.

Data Input Layer: The input data consist of process parameters X in the form of $M \times N \times d_x$, where M represents the number of samples, N denotes the time-series length, and dx is a feature dimension of 3.

Criterion Layer: The criterion layer comprises the process characteristic data, which serves as the output layer for the process parameters and the input layer for the weld quality. The data of the layer have dimension $M \times N \times d_z$, where d_z represents the maximum value that can be reached by the feature dimension in the process characteristic data, which is 9.

Self-attention Layer: The weights are calculated with the self-attention mechanism to focus adaptively on the key information in the sequence. MHSA-1 performs an attention operation on the process parameters and input process parameters $X \in R^{M \times N \times d_x}$. During the calculation, zero vectors are used to pad the feature dimensions that are smaller than dx. After the calculation, a new feature X' is obtained, with dimension $M \times N \times d_x$. Similarly, Z is calculated for the process characteristic data $Z \in R^{M \times N \times d_z}$.

LSTM Layer: LSTM neural networks are adept at processing time-series information and input vectors through recursive execution, depending on the past hidden states and current inputs. LSTM-1 is used to learn the time-series dependencies between the process parameters and the process characteristic data. The establishment of the nonlinear relationship can help achieve the precise control of the welding process. LSTM-2 is used to learn the time-series dependencies between the process characteristic data and the output to predict the final quality. The number of hidden nodes in both LSTM layers is set to L.

Output Layer: The output layer predicts the weld quality, with an output dimension $M \times d_y$, where d_y is 2.

Table 1. Layers of MHSA-LSTM-AHP Model.

No.	Layer	Data	Dimension
1	Data Input Layer	Process Parameters	$M \times N \times d_x$
2	MHSA-1	Weight Information	$M \times N \times d_{k1}$
3	LSTM-1	Hidden State	$N \times L$
4	Criterion Layer	Process characteristic data	$M \times N \times d_z$
5	MHSA-2	Weight Information	$M \times N \times d_{k2}$
6	LSTM-2	Hidden State	$N \times L$
7	Output Layer	Weld Quality	$M \times d_y$

4. RESULTS ANALYSIS AND DISCUSSION

In this study, the process data was collected on the welding process of a battery's terminal ears to validate the effectiveness of the model. The welding subjects are 40 layers of aluminum foil (single layer 12um) welded to a copper plate (0.8mm), with a welding head size of 16mm by 3mm. The set range of process parameters is: 50kPa < P < 150kPa, 16um < C1 < 28um, 1200w < W1 < 2500w, 0.2S < T1 < 0.8S, 0.1S < T2 < 0.4S, 4um < C2 < 9um, 200w < W2 < 400w, 0.05S < T3 < 0.2S. Ultrasonic welding was performed using different combinations of process parameters. The resistance tester Xinyang Electronics CXT2516 was used to measure the contact resistance of the joint, and the micro-computer-controlled universal testing machine MTS (CMT4304 model) was employed to measure the tensile shear strength of the joint, with a tensile speed set at 1mm/min. As the experiments in measuring the maximum tensile shear strength is destructive to the welded joints, 200 samples were collected and the k-fold cross-validation method was adopted to train and test the predictive model. Section 4.1 describes the training process of the model, Section 4.2 analyzes and validates the proposed model, and Section 4.3 compares the proposed model with other prediction models to evaluate its predictive performance.

4.1. MODEL TRAINING

The model training process is shown in Fig. 7 (see section: supplementary material). The main steps are described below.

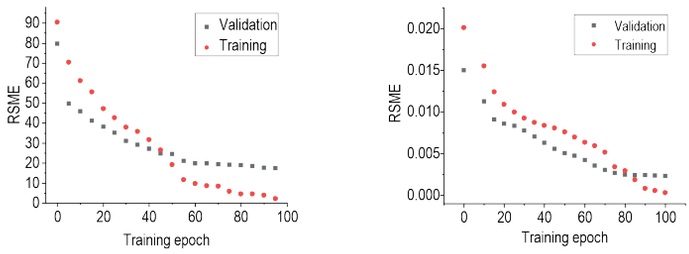
Data Processing: The sample data are normalized and divided into a training set, a validation set, and a test set.

Model Training: Weights are calculated layer by layer based on the AHP. The weights between the criterion layer and the output layer are calculated, followed by the weights between the input layer and the criterion layer. To mitigate the large fluctuations in the validation scores owing to the limited number of validation data points, k-fold cross-validation is employed for the model training.

K-fold Cross-validation: The dataset is partitioned into k parts of equal size. For each partition *i*, the model is trained on the remaining k-1 partitions and evaluated on partition *i*. The final score is the average of the score of all the partitions.

Model Loss Calculation: The RMSE is used as the loss function, which is calculated as follows:

$$RSME = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)^2} \tag{13}$$



(a) Tensile Shear Strength (b) Contact Resistance

Fig. 8. RMSE Curves of Training and Validation Sets.

Parameter Optimization: Gradient descent is used to adjust the model parameters until the loop termination condition is fulfilled. Once the process concludes, the optimal model parameters are saved for subsequent use.

This study compares the proposed model with the GA-BP Neural Network model [33-34] and the LSTM-AHP model without an attention mechanism to analyze the performance of the proposed MHSA-LSTM-AHP model in predicting the weld quality. The GA-BP Neural Network, LSTM-AHP, and MHSA-LSTM-AHP models are implemented by using the MATLAB neural network toolbox. The training and testing procedures are executed on an Nvidia GeForce RTX 3080 GPU platform with 12 GB of Video Random Access Memory (VRAM). A dropout structure, with a dropout rate of 0.2, is introduced into the MHSA-LSTM-AHP model to prevent overfitting, which means that 20% of the neurons are randomly dropped during the training. The dropout rate of the self-attention mechanism is set to 0.2, and the number of heads is set to 3. The LSTM neural network has 16 hidden nodes and 2 hidden layers. The adaptive moment estimation optimizer, with a learning rate of 0.001, is leveraged for the error backpropagation. The data samples are divided into training, validation, and test sets, with a ratio 76:19:5 and a batch size of 16 and 100 training epochs. After multiple adjustments, the k-value of the cross-validation is set to 5. The learning rate, number of iterations, and number of hidden layers in the two comparison models are the same as those in the proposed model.

4.2. QUALITY PREDICTION BASED ON MHSA-LSTM-AHP MODEL

The RMSE curves of the training and validation sets for the tensile strength and contact resistance prediction during the training of the proposed MHSA-LSTM-AHP model are depicted in Fig. 8.

Table 2. Predicted and Actual Tensile Strength and Contact Resistance Values.

No.	Tensile Shear Strength			Contact Resistance		
	Predicted (MPa)	Actual (MPa)	Relative Error(%)	Predicted (mΩ)	Actual (mΩ)	Relative Error(%)
1	10.271	10.813	5.01	0.052	0.051	1.96
2	6.854	7.292	6	0.057	0.06	5
3	13.458	13.333	0.94	0.04	0.042	4.76
4	12.083	12.333	2.03	0.049	0.049	0
5	9.437	9.125	3.42	0.052	0.053	1.89
6	9.895	9.625	2.81	0.053	0.051	3.926
7	11.271	11.709	3.74	0.044	0.046	4.345
8	9.146	9.334	2.01	0.046	0.049	6.12
9	8.688	8.355	3.99	0.058	0.056	3.57
10	11.625	11.376	2.19	0.052	0.055	5.45
Average Error			3.21			3.7

The results indicate that the model completes the error iteration convergence process during the tensile strength and contact resistance prediction training quickly and thus demonstrates excellent convergence performance. After 100 training epochs, the tensile strength and contact resistance RMSE value in the validation set is 15.26 and 0.0027, respectively, which show small errors compared with the actual values. Although the convergence speed slows in the later stages of the training of the validation set, it continues to decrease and gradually stabilizes, which indicates that no overfitting occurs and confirms the effectiveness of the model training.

Randomly selected 10 samples from the 200 samples were used as a test set to evaluate the predictive ability of the model. Table 2 presents the actual measured values of the maximum tensile shear strength and contact resistance of the welding samples under different combinations of process parameters, as well as the predicted values calculated by the model. The comparison of the predicted and actual values is illustrated in Fig. 9.

Table 2 and Fig. 9 demonstrate that the MHSA-LSTM-AHP prediction model achieves an average relative percentage error of 3.21% and 3.7% for tensile strength and contact resistance, respectively. The data indicate that the predictions are close to the actual measurements, and the model has high predictive and generalization capabilities. However, the prediction error for contact resistance is slightly higher than that for tensile strength, which can be attributed primarily to measurement errors. Contact resistance is small and requires high measurement accuracy, making it susceptible to factors such as the condition of the weld contact surface and the environmental interferences, which can reduce the accuracy of the model training.

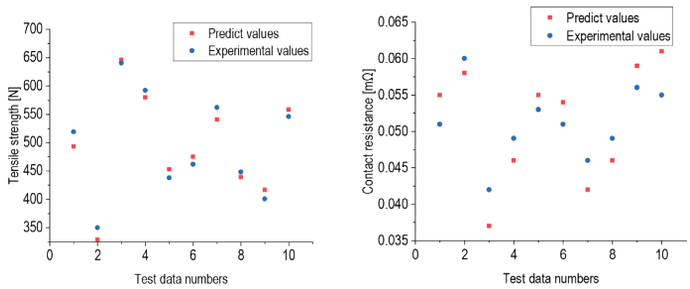
4.3. COMPARISON OF PREDICTION PERFORMANCE

To verify whether the introduction of MHSA into the LSTM-AHP model can effectively improve its predictive ability for weld quality and analyze the performance of the proposed model, this study compares the quality predictions of three trained models, namely, the MHSA-LSTM-AHP, LSTM-AHP, and GA-BP Neural Network models, by using the same test set. The tensile strength and contact resistance prediction results are shown in Fig. 10 and Fig. 11 (see section: supplementary material).

The comparisons show that the tensile strength and contact resistance values predicted by the MHSA-LSTM-AHP model are in best agreement with the experimental values, which demonstrate the model's superior predictive ability. Although the prediction error of the LSTM-AHP neural network model increases to a certain extent, compared with the GA-BP Neural Network model, the LSTM-AHP neural network model demonstrates higher accuracy. In addition, the three models exhibit a more pronounced degree of deviation when predicting contact resistance compared with tensile strength. This phenomenon lends support to the hypothesis that the relatively large error in the resistance measurements is attributable to measurement issues, rather than inherent flaws in the models.

To accurately evaluate the models, this study selects the mean absolute error(MAE), RMSE, and mean absolute percentage error (MAPE) to analyze the results of the validation set.

As shown in Table 3, in predicting tensile strength, the MAE, RMSE, and MAPE of the MHSA-LSTM-AHP model are 0.31, 0.336, and 0.032, respectively, which are lower than those of the LSTM-AHP and GA-BP Neural Network models. Similarly, the errors are significantly reduced in the contact resistance prediction. Moreover, a substantial reduction can be seen in the errors when the model predicts resistance. The result demonstrates that the MHSA-LSTM-AHP model has superior prediction performance.



(a)Tensile Shear Strength *(b) Contact Resistance*

Fig. 9. Comparison of Predicted and Actual Values Based on MHSA -LSTM-AHP Model.

In Comparison to the performance of the LSTM-AHP model, the MAE, RMSE, and MAPE of the proposed model decrease by 17%, 13.4%, and 16.3%, respectively. Moreover, in the resistance prediction, the values decrease markedly by 24%, 21.5%, and 24.7%. The results indicate that MHSA introduced into the LSTM-AHP model can considerably enhance the model's predictive ability.

The errors of the GA-BP Neural Network, LSTM-AHP, and MHSA-LSTM-AHP models increase in the resistance prediction compared with those in the tensile strength prediction. However, the increase in the errors of the GA-BP Neural Network model is substantial. This result suggests that the LSTM-AHP and MHSA-LSTM-AHP models, which are capable of learning temporal dependencies, can extract substantial information and demonstrate superior robustness. The prediction error of contact resistance is decreased with incorporating MHSA, as this mechanism utilizes multiple independent attention heads to learn diverse representations of data from different subspaces. This allows the model to further learn key information between the process parameters and the quality indicators from more dimensions, thereby making it more robust to data noise or partial feature absence.

To better characterize the coupling relationship between process parameters and quality in UMW process, this study establish an AHP-based hierarchical relationship among process parameters, process characteristics, and welding quality. The LSTM units are utilized to learn the temporal relationships between parameters and the final welding quality indicators throughout the welding process. Moreover, the model integrates the MHSA mechanism to enhance its predictive capabilities across various dimensions. To scientifically validate the proposed model, this study conducted training, learning, and comparative analysis. The results indicate that our model possesses relatively high predictive accuracy.

5. CONCLUSION

To investigate the relationship between the process parameters in each stage and the final joint quality in the UMW process, this study conducts a detailed analysis of the process parameters

Table 3. Performance of Each Model.

	Algorithm Type	MAE	RMSE	MAPE
Tensile Strength	MHSA-LSTM-AHP	0.31	0.336	0.0321
	LSTM-AHP	0.38	0.389	0.0384
	GA-BP Neural Network	0.52	0.543	0.0533
Contact Resistance	MHSA-LSTM-AHP	0.0019	0.0021	0.0370
	LSTM-AHP	0.0025	0.0027	0.0492
	GA-BP Neural Network	0.0034	0.0036	0.0679

and process characteristic data during the welding process. By leveraging the AHP, this study clearly delineates the hierarchical and sequential structures between each process parameter and the weld quality during the welding process. By using the LSTM and MHSA methods, this study establishes an MHSA-LSTM-AHP-based quality prediction model to accurately predict the weld quality of UMW. Furthermore, this study conducts a comparative analysis between the proposed model and other quality prediction models and draws the following conclusions:

(1) Compared with traditional prediction models, the proposed MHSA-LSTM-AHP model achieves smaller errors when predicting tensile strength and contact resistance and thus demonstrates higher prediction accuracy.

(2) The integration of MHSA into the LSTM-AHP model can substantially improve the model's weld quality prediction accuracy.

(3) By utilizing the memory units of LSTM and the multi-head mechanism of MHSA to capture the dependency relationships between process parameters and quality indicators from multiple dimensions, the proposed model demonstrates strong robustness during the prediction of contact resistance, which is a small numerical value and is deeply influenced by the condition of the weld contact surface.

The research results contribute to UMW systems that can achieve the online prediction and evaluation of weld quality, which can be used to provide a reference for the optimization of UMW process parameters in the production of key electronic components.

REFERENCES

[1] Das A, Li D, Williams D, et al. "Joining technologies for automotive battery systems manufacturing". World Electric Vehicle Journal. July 2018. Vol.9-2. p.22. DOI: https://doi.org/10.3390/wevj9020022
[2] Arimoto K, Sasaki T, Doi Y, et al. "Ultrasonic bonding of multi-layered foil using a cylindrical surface tool". Metals. April 2019. Vol.9-5. p.505. DOI: https://doi.org/10.3390/met9050505
[3] Shin S, Nam S, Yu J, et al. "Ultrasonic metal welding of multilayered copper foils to nickel-plated copper sheet in lithium-ion battery cell". Metals. July 2021. Vol.11-8. p.1195. DOI: https://doi.org/10.3390/met11081195
[4] Shi X, Li L, Yu S, et al. "Process monitoring in ultrasonic metal welding of lithium batteries by power signals". Journal of Manufacturing Science and Engineering. May 2022. Vol.144-5. p.051007. DOI: https://doi.org/10.1115/1.4052704
[5] Lee SS, Kim TH, Hu SJ, et al. "Characterization of joint quality in ultrasonic welding of battery tabs". Journal of Manufacturing Science and Engineering. June 2012. Vol.135-2. p.249-261. DOI: https://doi.org/10.1115/1.4023364
[6] SAE. "USCAR38-2: Performance specification for ultrasonically welded wire terminations". SAE International. July 2023. DOI: https://www.sae.org/standards/content/uscar38-2/
[7] Nong L, Shao C, Kim TH, et al. "Improving process robustness in ultrasonic metal welding of lithium-ion batteries". Journal of Manufacturing Systems. July 2018. Vol.48. p.45-54. DOI: https://doi.org/10.1016/j.jmsy.2018.04.014
[8] De Leon M, Shin HS. "Prediction of optimum welding parameters for weld-quality characterization in dissimilar ultrasonic-welded Al-to-Cu tabs for Li-ion batteries". Metals and Materials International. April 2023. Vol.29-4. p.1079-1094. DOI: https://doi.org/10.1007/s12540-022-01271-z
[9] Feng M, Wang Z, Meng D, et al. "A review of quality monitoring for ultrasonic metal welding". Materials Science and Technology. Janurary 2024. Vol.40-1. p.3-25. DOI: https://doi.org/10.1177/02670836231215651
[10] Zhao D, Jiang C, Zhao K. "Ultrasonic welding of AZ31B magnesium alloy and pure copper: Microstructure, mechanical properties and finite element analysis". Journal of Materials Research and Technology. March 2023. Vol.23. p.1273-1284. DOI: https://doi.org/10.1016/j.jmrt.2023.01.095
[11] Ma Z, Luo C, Zhang Y. "Prediction of the interface deformation of ultrasonic spot welding of multilayer metal foils considering energy gradient". Journal of Manufacturing Science and Engineering. May 2022. Vol.144-5. p.051011. DOI: https://doi.org/10.1115/1.4053924
[12] Ebrahimi A, Bayat M, Norouzi E. "Measurement of residual stress using the ultrasonic method in aluminum welding: FE analysis and experimental study". Russian Journal of Nondestructive Testing. August 2021. Vol.57-8. p.669-678. DOI: https://doi.org/10.1134/S1061830921080027
[13] Shen N, Samanta A, Cai WW, et al. "3D finite element model of dynamic material behaviors for multilayer ultrasonic metal welding". Journal of Manufacturing Processes. February 2021. Vol.62. p.302-312. DOI: https://doi.org/10.1016/j.jmapro.2020.12.039
[14] Li H, Zhang CX, Zhou K, et al. "Effect of vibration amplitude on ultrasonic welding of Cu/Al". Transactions of the China Welding Institution. July 2023. Vol.44-7. p.40-47. DOI: https://doi.org/10.12073/j.hjxb.20220815001
[15] Yao Z, Liu K, Ding R, et al. "Optimization of ultrasonic welding parameters based on response surface methodology". Journal of Physics: Conference Series. April 2023. Vol.2474-1. p.012029. DOI: https:// doi.org/10.1088/1742-6596/2474/1/012029
[16] Al Sarraf Z. "Prediction of weld strength in power ultrasonic spot welding process using artificial neural network (ANN) and back propagation method". Journal of Manufacturing Engineering. December 2022. Vol.17-4. p.119-126. DOI: https://doi.org/10.37255/jme.v17i4pp119-126
[17] Mongan PG, Hinchy EP, O'Dowd NP, et al. "Quality prediction of ultrasonically welded joints using a hybrid machine learning model". Journal of Manufacturing Processes. November 2021. Vol.71. p.571-579. DOI: https://doi.org/10.1016/j.jmapro.2021.09.044
[18] Lee SS, Shao C, Kim TH, et al. "Characterization of ultrasonic metal welding by correlating online sensor signals with weld attributes". Journal of Manufacturing Science and Engineering. October 2014. Vol.136-5. p.051019. DOI: https://doi.org/10.1115/MSEC2014-4139.
[19] Choi H, Shin S, Kim DY, et al. "Investigation on quality prediction algorithm in ultrasonic metal welding for multilayered Cu for battery cells". IEEE Access. 2023. Vol.11. p.146313-146321. DOI: https://doi.org/10.1109/ACCESS.2023.3344160
[20] Guo W, Shao C, Kim TH, et al. "Online process monitoring with near-zero misdetection for ultrasonic welding of lithium-ion batteries: An integration of univariate and multivariate methods". Journal of Manufacturing Systems. January 2016. Vol.38. p.141-150. DOI: https://doi.org/10.1016/j.jmsy.2016.01.001
[21] Hong S, Cho J, Bang H. "Quality classification of multi-layer copper foil stacks welds for ultrasonic welding using CNN and RNN". Journal of Manufacturing Systems. October 2024. Vol.42-5. p.514-520. DOI: https://doi.org/10.5781/JWJ.2024.42.5.8
[22] Samir S, Dave K, Badheka V, et al. "Optimization of process parameters of ultrasonic metal welding for multi-layers foil of AL8011 material". Welding International. March 2023. Vol. 37-3. p.119-127. DOI: https://doi.org/10.3390/polym14122388
[23] Müller FW, Mirz C, Weil S, et al. "Weld quality characterization by vibration analysis for ultrasonic metal welding processes". Journal of Advanced Joining Processes. November 2023. Vol.8. p.100149. DOI: https://doi.org/10.1016/j.jajp.2023.100149
[24] Müller FW, Chen CY, Schiebahn A, et al. "Application of electrical power measurements for process monitoring in ultrasonic metal welding". Welding in the World. December 2022. Vol.67-2. p.395-415. DOI: https://doi.org/10.1007/s40194-022-01428-9
[25] Schwarz EB, Bleier F, Guenter F, et al. "Improving process monitoring of ultrasonic metal welding using classical machine learning methods and process-informed time series evaluation". Journal of Manufacturing Processes. May 2022. Vol.77. p.54-62. DOI: https://doi.org/10.1016/j.jmapro.2022.02.057
[26] Chen H, Dong X, Kong Y, et al. "Online prediction of mechanical and electrical quality in ultrasonic metal welding using time series generation and deep learning". Engineering Failure Analysis. June 2024. Vol.160. p.108162. DOI: https://doi.org/10.1016/j.engfailanal.2024.108162
[27] Wu Y, Meng Y, Shao C. "End-to-end online quality prediction for ultrasonic metal welding using sensor fusion and deep learning". Journal of Manufacturing Processes. November 2022. Vol.83. p.685-694. DOI: https://doi.org/10.1016/j.jmapro.2022.09.011
[28] Munier N, Hontoria E. "Uses and limitations of the AHP method: A non-mathematical and rational analysis". Cham, Switzerland: Springer International Publishing. 2021. DOI: https://doi.org/10.1007/978-3-030-60392-2
[29] Li Y, Qin H, Zhang J, Su H, Li G, Bai S. "Cloud Model for Security State Recognition Based on Factor Space".IEEE Sensors Journal. November 2021.Vol. 21- 22. p. 25429-25436. DOI:https://doi.org/10.1109/ JSEN.2021.3098679
[30] Alzubaidi L, Zhang J, Humaidi AJ, et al. "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions". Journal of Big Data. March 2021. Vol.8-1. p.53. DOI: https://doi.org/10.1186/s40537-021-00444-8
[31] Yu Y, Si X, Hu C, et al. "A review of recurrent neural networks: LSTM cells and network architectures". Neural Computation. July 2019. Vol.31-7. p.1235-1270. DOI: https://doi.org/10.1162/neco_a_01199
[32] Yuan M, Jia W, Luo X, et al. "MHSA: A multi-scale hypergraph network for mild cognitive impairment detection via synchronous and attentive fusion". In: Proceedings of 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Lisbon, Portugal: IEEE, December 2024. p.2808-2815. DOI: https://doi.org/10.1109/BIBM62325.2024.10822410
[33] Jin L, Duan J, Fan T, et al. "Using GA-BP coupling algorithm to predict the high-performance concrete mechanical property". KSCE Journal of Civil Engineering. February 2023. Vol.27-2. p.684-697. DOI: https://doi.org/10.1007/s12205-022-0912-9
[34] Yang L,Zhong J H,Zhang Y, Bai S C, Li G S, Yang Y, Zhang J. "An Improving Faster-RCNN With Multi-Attention ResNet for Small Target Detection in Intelligent Autonomous Transport With 6G".IEEE Transactions on Intelligent Transportation Systems.July 2023. Vol.24-7.p.7717-7725.DOI:https://doi. org/10.1109/TITS.2022.3193909

SUPPLEMENTARY MATERIAL

https://www.revistadyna.com/documentos/pdfs_adic/11433-1_en.pdf