

Online quality inspection of resistance spot welding for automotive production lines

Wei Dai^{a,b}, Dayong Li^{a,b}, Yongjia Zheng^a, Dong Wang^a, Ding Tang^{a,*}, Huamiao Wang^{a,b}, Yinghong Peng^a

^a State Key Laboratory of Mechanical Systems and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China

^b Materials Genome Initiative Center, Shanghai Jiao Tong University, 200240, China



ARTICLE INFO

Keywords:

Quality inspection
Process stability
Deep learning
Low-rank and sparse decomposition
Channel attention mechanism

ABSTRACT

Reliable quality control of resistance spot welding (RSW) is a long-standing challenge, due to random disturbance on automotive production lines. In this paper, a quality evaluation framework is proposed based on dynamic resistance (DR) signals, aiming to accurately predict welding quality. The proposed framework integrates welding process stability with deep learning models. Given the uniform variation pattern of each weld with the same schedule, process stability can be determined based on the reference curve constructed by the low-rank and sparse decomposition method. Subsequently, a one-dimensional convolutional neural network (1DCNN) with channel attention mechanism is developed to further predict welding quality, which can perform channel-wise feature recalibration to enhance the classification performance. Extensive experiments substantiate that the proposed network yields a remarkable classification performance compared with typical algorithms on several RSW datasets collected on an actual production line. This study provides a valuable reference to achieve an intelligent online quality inspection system in the automotive manufacturing industry.

1. Introduction

Resistance spot welding (RSW) is an ideal choice for automotive manufacturing, due to its low cost, high operating speeds, and suitability for automation [1]. With the development of the industry, the quality of spot welds has become more demanding in various applications. However, RSW is a highly non-linear, multivariate coupling process with many uncertainties, making it difficult to predict quality accurately with physical models [2]. In high-volume and fast-paced production lines, it is usually accompanied by disturbing conditions such as electrode wear, assembly gaps, and workpiece surface contamination [3,4]. These random uncertainties increase the instability of the welding process, leading to welding quality problems such as cold welds and expulsion, which reduces the reliability of the car body. Traditionally, manual sampling inspection is carried out regularly in the production site to ensure welding quality. This inspection method is susceptible to randomness and cannot fully guarantee the quality of spot welds, which is not only laborious but also leaves safety concerns. More importantly, these inspection methods tend to be predominantly offline, which makes it impossible to use this information to control the welding process.

Therefore, it is urgent to find an effective and accurate RSW inspection technology for online full inspection of welding quality at the production site, thus realizing a high-quality intelligent automobile production line.

The advent of low-cost sensors and advances in wireless communications have allowed body manufacturers to access vast amounts of data for the manufacturing process [5–7]. The accumulation of process data provides the opportunity to better monitor the welding process. Developments in emerging approaches such as the Internet of Things (IoT), machine learning, and deep learning have also reinvigorated research in this area [8–10]. Great efforts had been put into the research of online monitoring for welding quality during the RSW process. Among them, various process signals had been used to study welding quality, e.g., dynamic resistance (DR) [11,12], electrode force [13,14], electrode displacement [15,16], and acoustic emission [17,18].

Current research on process signal-based spot welds quality prediction is often associated with machine learning models [19]. The routine process of quality prediction by machine learning models is as follows: firstly, several features related to welding quality are extracted manually from the process signals. Subsequently, an appropriate machine learning

* Corresponding author.

E-mail address: tangding@sjtu.edu.cn (D. Tang).

algorithm is selected to model the relationship between the features and welding quality. Finally, model parameters are trained by the dataset obtained in the laboratory and an online quality evaluation model is established. Specifically, twenty relevant features were extracted from DR signals, and a mathematical model for weld quality prediction was developed based on a stepwise regression approach [20]. DR signal features were utilized to estimate welding quality by multiple linear regression and backpropagation neural network (BPNN) models [21]. BPNN and probabilistic neural network (PNN) models were proposed for establishing the relationship between extracting features from DR signals and weld quality [22]. An online approach for welding monitoring was proposed to effectively determine spot weld quality, which combines DR signal features and random forest algorithm [23]. Multidimensional features calculated from displacement signals were drawn as Chernoff faces to characterize welding quality [24,25]. A combination of welding process feature analysis extraction, principal component analysis, K-means clustering algorithm, and BPNN was proposed for the prediction of the quality [26].

Although considerable progress has been made in quality prediction, there are still some flaws in the traditional welding monitoring technology. Its detection accuracy and reliability are not yet able to meet the actual production requirements, and there is still a large gap between the goal of achieving intelligent welding quality monitoring. Firstly, the existing machine learning-based methods lack universality and generalization. Feature engineering of traditional machine learning models requires rich a priori knowledge and the risk of collapsing useful information. As a result, the performance is highly reliant on whether extracted features are correlated with welding quality. In addition, most of the research results are obtained from offline experiments under laboratory conditions based on a few types of material stack-up combination and working conditions [20,22,23], all of which have limitations in practical applications. Due to the complexity and uncertainties of actual welding conditions, it is difficult to extend these quality inspection methods to practical automotive production lines. Secondly, there is a class imbalance problem in the collected spot welding datasets. The availability of large-scale annotated and class-balanced datasets is of great importance for learning-based models. However, it is very expensive and time-consuming to collect defective spot welding data on the production line. This results in serious class-imbalanced of the collected spot welding dataset, which will deteriorate the performance of the learning-based model. Finally, an entirely learning-based approach lacks physical significance and interpretability. Most scholars used machine learning models directly to establish the relationship between the extracted features and welding quality. Such methods are only concerned with the input and output, the internal mechanism is equivalent to a black box [2,19]. This leads to a lack of interpretability of the built models, which affects the reliability of the quality inspection approach.

Recently, deep learning models have achieved breakthrough improvement in automatically extracting highly discriminative features for time series classification [27,28]. Deep learning models enable automatic feature extraction of time series, which can substitute the requirement of the additional handcrafted feature extraction step mentioned above. Among deep learning model architectures, one-dimensional convolutional neural network (1DCNN) has shown promising performances on time series classification [29,30]. In automotive manufacturing companies, sensors mounted on welding robots sample time-series signals all the time, thus generating a huge amount of welding process data. Traditional machine learning methods are difficult to deal with big data effectively, while deep learning is the most effective method to deal with big data at present. Theoretically, the features extracted by deep learning models become more effective with the increase of data volume, and the prediction accuracy increases. Deep learning methods have been successfully applied in welding fields and have shown promising performances in this field [31–34]. For example, the bayesian regularized neural network and CNN models were trained

to predict weld quality classifications in ultrasonic welding [35,36]. A long short-term memory (LSTM) recurrent neural network was proposed to monitor the ultrasonic welding process and predict quality class from continuous signals [37]. The online detection of porosity defects during aluminum alloys laser welding was completed by a CNN-based deep learning model [38]. However, there is little research on spot welding quality prediction based on process signals with deep learning methods.

The quality evaluation algorithm based entirely on deep learning models lacks physical significance, and it depends on the quality and quantity of training samples. In automotive production sites, the collected signals are susceptible to unpredictable disturbances and anomalies, and these data can affect the training of learning models, thereby deteriorating performances of quality evaluation algorithms. Therefore, an RSW quality inspection approach is proposed in the present paper, which integrates welding process stability and deep learning models. The welding process stability is determined by constructing a reference curve and similarity measure for each spot weld. This method is conducive to screening out abnormal signals in mass production. Spot welds at the same position on different vehicles have the same welding parameters, and the collected DR signal is theoretically similar. Therefore, the measured signal matrix has low-rank components, which can be used to construct a reference curve. Here, the low-rank and sparse decomposition method [39] is utilized to obtain low-rank components of the measured signal matrix. This strategy has a clear physical meaning and can accurately characterize the consistency of welding quality. Subsequently, inspired by attention mechanisms [40–42], 1DCNN model combines channel attention mechanism and residual connection is proposed for quality prediction. The channel attention mechanism considers the weight of different feature maps to enhance classification performance, while the residual connection structure can promote feature reuse and reduce redundancy. The purpose of the present work is mainly to detect defective spot welds caused by uncertain factors in a plant environment. Extensive experimental studies are also performed to demonstrate the accuracy and generalization of the proposed prediction approach. In conclusion, the main contributions of this paper are as follows:

- (1) Based on dynamic resistance signals, an RSW quality inspection approach is developed for automotive production lines. The reference curve for each spot weld with the same schedule is constructed by low-rank and sparse decomposition method and welding stability based on the reference curve used as a priori condition.
- (2) A deep network combining channel attention mechanism and residual connection is proposed to further predict the welding quality. The proposed network is quantitatively validated on spot welds datasets under different welding conditions, and promising results are obtained compared with the typical deep learning models.
- (3) The developed quality inspection approach requires little prior expertise on resistance spot welding and feature engineering, which facilitates its extension to actual automobile production lines.

The rest of this paper is organized as follows. Section 2 introduces welding data acquisition and preprocessing. The proposed quality inspection approach is described elaborately in Section 3. In Section 4, the performance of the quality inspection approach is comprehensively validated on the different spot welding datasets. Finally, the conclusion will be drawn in Section 5.

2. Data acquisition and preprocessing

The demonstrated local automobile production lines employ robotic spot welding processes for manufacturing. The welding is carried out on Fanuc robotic welding system equipped with servo welding guns and

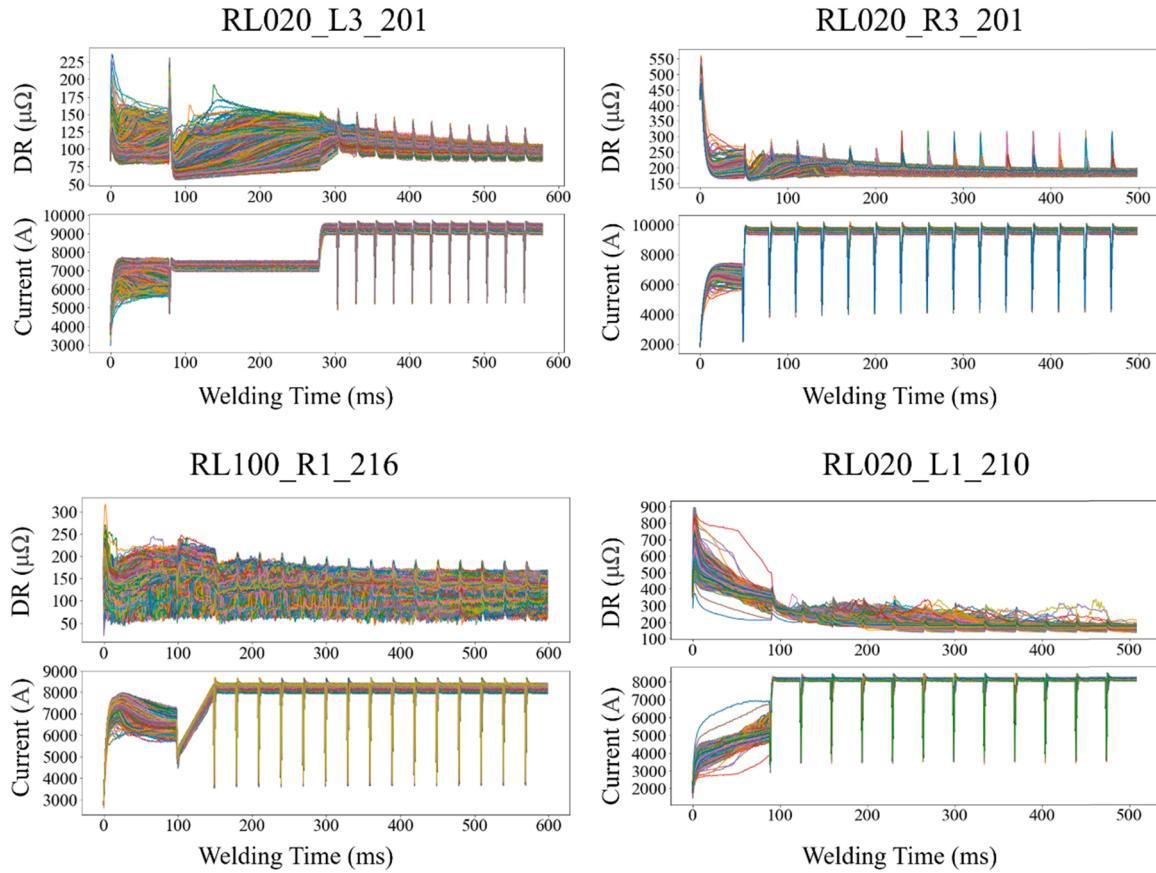


Fig. 1. Illustrations of collected DR and welding current signals for different spot welds. Note: RL090_L2_201 means the process signals collected from the L2 welding torch at 090 station, and the schedule number is 201. Same for RL100_R1_216, RL020_L3_201, and RL020_R3_201.

MEDAR/WTC medium frequency direct current (MFDC) controllers. The workpiece is assembled with a lap structure of both two and three layers of plates, and the thicknesses of the sheet are 0.65 mm, 0.70 mm, 1.0 mm, and so on. Welded materials include cold-rolled steels, hot press forming steels, dual-phase steels, and the surface coating includes two types of galvanized and aluminum-silicon coating, which are commonly used in automotive body-in-white.

A typical RSW operation is controlled by a welding schedule, whose time steps are controlled by a welding controller. The welding parameters that will be used for each spot are programmed into the weld-timer controller associated with the gun which is used for that particular spot. With the simultaneous collection of electrode voltage and welding current signals by the monitoring system on the production lines, DR signals are further calculated by the software embedded. The above process signals are stored in the database of each spot of each autobody against the timestamp when the spot is made.

2.1. Production line welding data

A mass of RSW process signals are collected through collaboration with a local automotive production company. In this work, DR and welding current signals are measured and analyzed for each spot weld, as shown in Fig. 1. In high-volume automated production, spot welds at the same position on different vehicles have the same welding conditions and schedule. According to Fig. 1, a uniform trend of variation in DR signals of the same welding schedule can be observed, but it showed a certain dispersion. The resistance varies because of random factors on production lines, like the thickness of the coating, dust oil, improper weld tip contact, electrode wear, etc. Moreover, due to the difference in material stack-up combination and welding parameter settings, there are significant differences in DR signals under different schedules.

On the automotive production line, multi-pulse current welders with a preheating current are used, due to the high strength of body materials. The basic welding schedule consists of several constant current pulses with certain pulse duration, and cooling time between each pulse. The collected DR signals have many sudden changes due to the removal of the cooling section. In order to improve the quality evaluation performance, it is necessary to preprocess original signals. The spikes can be removed according to the current pulse time. In addition, there is some high-frequency noise in the main welding section, and it is sufficient to reduce the noise with a low-pass filter. The raw and preprocessed DR signals with different schedules are displayed in Fig. 2.

Quality assessment of spot welds in the automotive industry is typically based on cost- and time-consuming off-line manual tests that are unfeasible on the full production, especially on large scale. In this work, the quality of spot welds can be classified as “normal”, “burn through”, and “cold weld”, wherein “normal” is a spot with an acceptable nugget diameter. During the RSW process, it is often the phenomenon that molten metal is forcefully ejected from the welds, which refers to the expulsion. It does not always destroy the weld strength and small expulsion with acceptable nugget size can be considered normal welds on a practical production line. However, severe expulsion may eject a large amount of material to create a through-hole in the workpiece, commonly termed “burn through”. Burn through occurs leaving a hole completely through the welds, which need to be avoided. Cold weld indicates that the nugget diameter of the spot weld is smaller than the specified minimum diameter, or even no nugget. Cold welds seriously affect the safety performance of the vehicles and is also a defect that needs to be focused on in the production lines. The appearance quality of spot welds is marked by a welding expert engineer's visual inspection. Meanwhile, ultrasound testing is adopted on randomly sampled workpieces at the production site at regular intervals to determine the

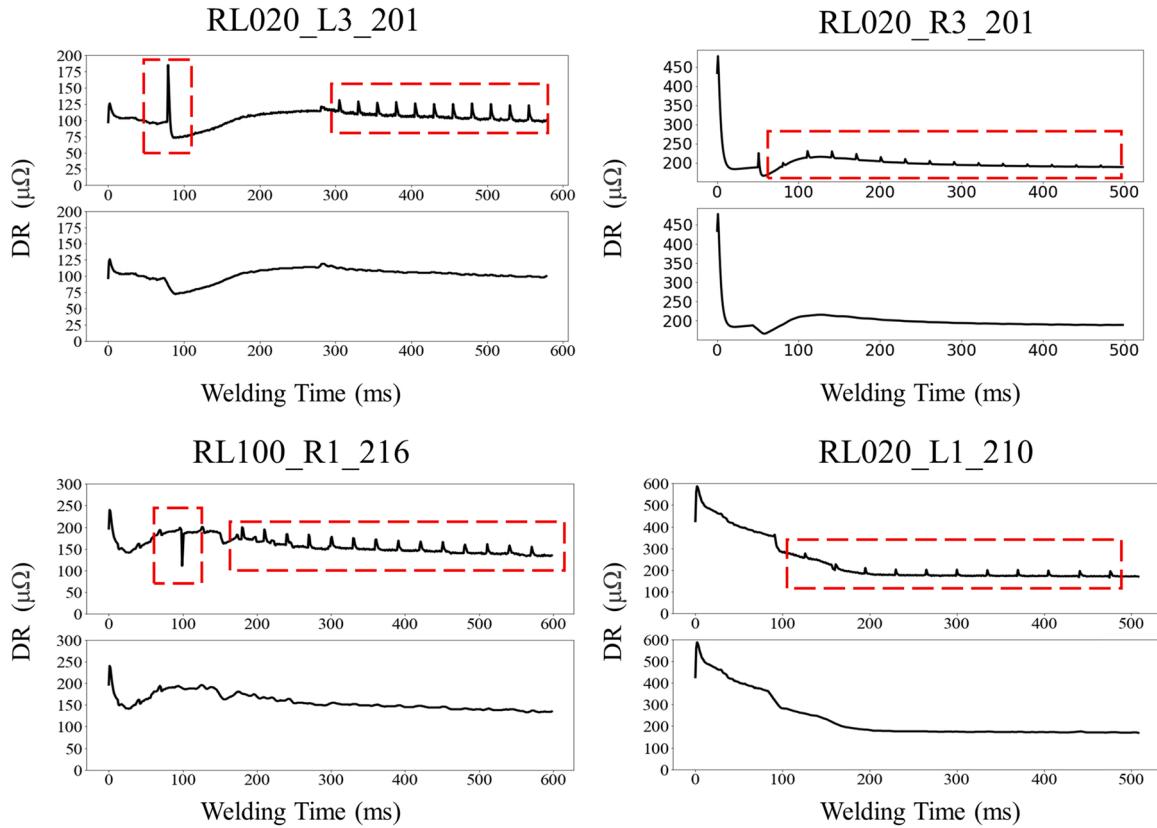


Fig. 2. Examples of preprocessing of DR signals with different schedules. Note: The spikes in the red box are to be preprocessed.

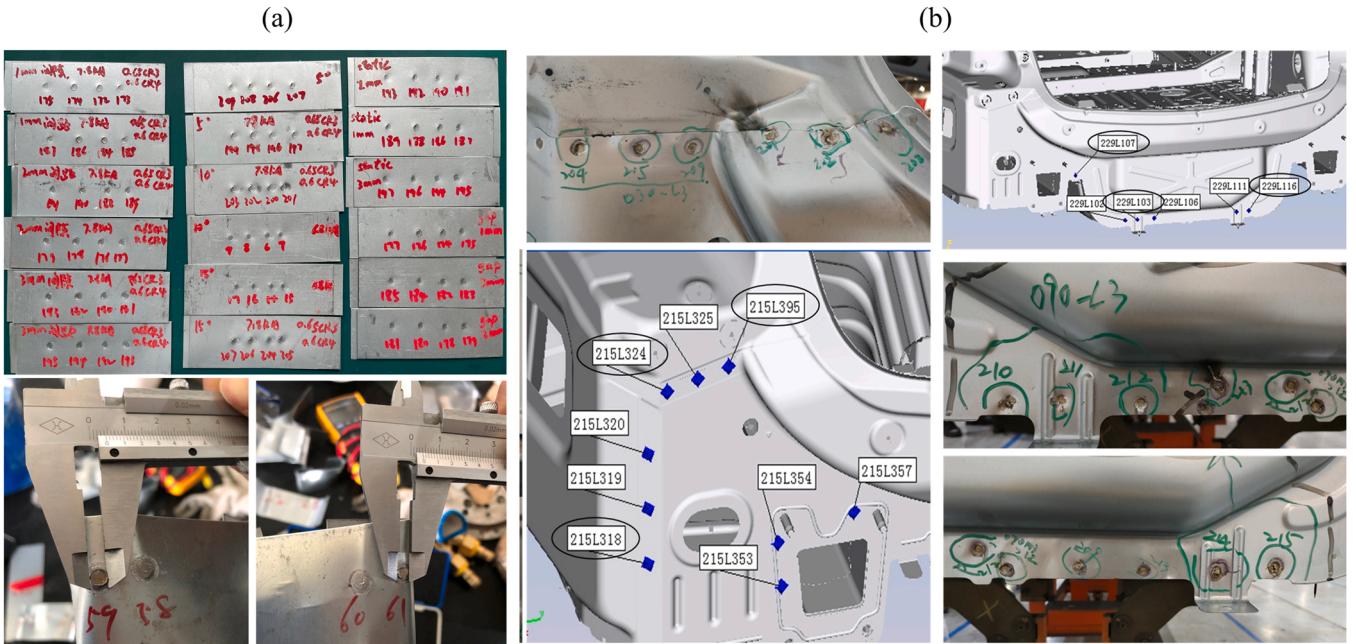


Fig. 3. Abnormal welding conditions experiments. (a) specimens; (b) vehicles.

welding quality. Due to the low defect rate, normal welds account for the vast majority, while the defective spot is extremely small on production lines. Thus, the ratio of defectives to non-defective spots is small, making the dataset obviously imbalanced.

2.2. Experimental welding data

It is very expensive and time-consuming to collect defective spot welding data on the production line. Therefore, in order to produce the typical welding defect, the same welding schedules and material stack-up combination as the production line are used to conduct welding

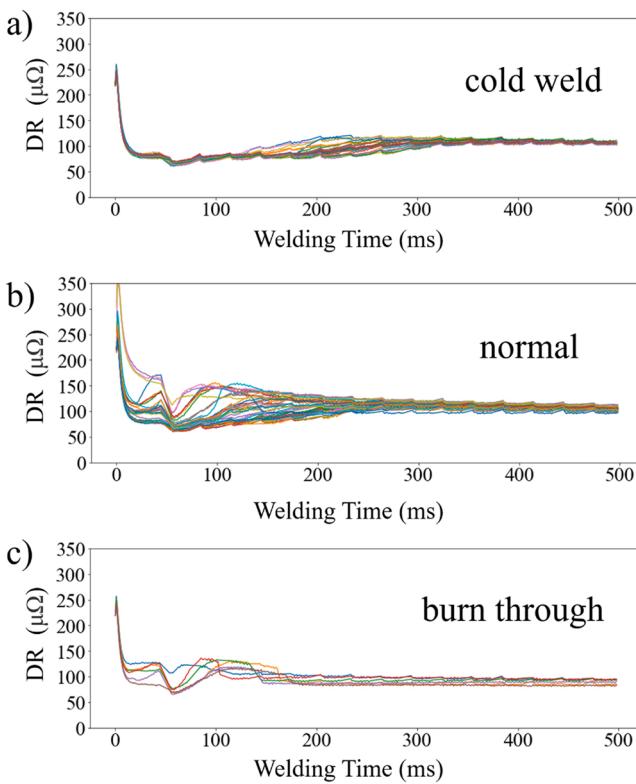


Fig. 4. Comparison of DR signals measured under three weld quality levels in abnormal experiments. a) cold weld; b) normal; c) burn through.

experiments. Various abnormal welding conditions that may occur in automotive manufacturing are adopted in the experiments, such as current shunt, sheet gap, sheet deflection angle, and so on. Commonly, these abnormal welding conditions can affect the dynamic distribution of the welding current density at each stage of the RSW, which further impacts weld nugget formation and growth. For example, the shunting current in RSW will deliver an apparent decrease in nugget diameter in consecutive welds on the same sheet, due to declined welding current [43]. The existence of the sheet gap condition results in directional weld nuggets in the lengthwise and widthwise dimensions respectively, which decreased the tensile-shear strength of the welds [44]. The deflection of the sheet relative to the weld gun causes asymmetric electrode-sheet contact. The asymmetrical heat input produced variable weld structure and performance, which affects the overall mechanical properties of welds [45].

Abnormal welding conditions experiments are conducted in both specimens and vehicles, as displayed in Fig. 3. The peel test is one of the most common destructive tests to check the quality of nuggets. The welded sheets get pulled apart and the corresponding weld nugget size is measured. The welding quality is also determined by ultrasound tests in the vehicle experiment. Finally, the nugget quality check label and the corresponding current and DR signals are stored manually in offline files.

Fig. 4 illustrates some DR signals measured under abnormal welding conditions with the same welding schedule in experiments. By combining physical characteristics of the dynamic resistance signal, notable morphological differences can be observed from different quality levels. For normal welds, their tendencies precisely resemble the typical DR curve of mild steel in RSW. However, cold welds do not obtain an apparent peak in the main welding section, suggesting minor local melting and inadequate nugget formed. When burn through phenomenon occurs, the ejection of molten metals induces a substantial reduction in the contact area. Hence, a remarkable drop in amplitude of the DR signal is indicated in Fig. 4. Furthermore, some signals do not

show such a law, and these signals are difficult to judge with human experience. Therefore, further intelligent quality prediction algorithms are desired.

3. Proposed quality inspection approach

The procedure for the proposed spot welds quality evaluation approach is illustrated in Fig. 5. It contains signal acquisition and pre-processing, welding process stability judgment, and deep learning-based classifier construction. First, through networking of all welding machines on the production line, welding current and DR signals of each spot can be collected and preprocessed, and a large database of spot welding quality information can be established. It can be found that there is a uniform variation trend in the DR signals with the same welding schedule. In response to the phenomenon, this paper introduces welding process stability based on the reference curve as an evaluation index for quality consistency. Next, the low-rank and sparse decomposition method is used to construct reference curves, based on a series of DR signals of normal spot welds. The data filtered by the reference curve will not contain abnormal signals caused by production disturbances. Its high data quality is conducive to the training of deep learning models. Therefore, a 1DCNN model that combines channel attention mechanisms with residual connection is finally established and trained on the filtered RSW datasets. In practical application, the reference curves and deep network weights obtained by offline training will be stored in the database. When an online quality inspection of RSW is carried out, the collected DR signal will be preprocessed first. Subsequently, welding process stability based on the reference curve is used to preliminarily determine the spot quality. In this work, spot welds that are highly consistent with the reference curve are considered as good quality. Specifically, the algorithm will call the corresponding reference curve from the database and calculate the similarity with the collected DR signal. If it is less than the set threshold, the spot welding quality is good; otherwise, the deep network model will make further quality judgment. The spot quality can be predicted by loading the well-trained network weights.

3.1. Welding process stability judgment

Through analysis of on-site body spot welding monitoring data, spot welds at the same location on different vehicles have similar welding conditions, including material stack-up combinations, welding current and welding time, etc. Therefore, the measured DR signal under the same welding specification will have a high degree of consistency. In addition to the spot welding quality, welding consistency and stability are also common concerns of automotive production companies. There is a uniform trend of variation with the same schedule in DR signals, as shown in Fig. 1. The smaller the dispersion, the higher the stability of the welding process. The difference between the measured DR signal and the uniform variation pattern can be used as an indicator of welding process stability. Moreover, the results of the online stability evaluation can be used as an a priori condition for other quality evaluation methods. The uniform variation pattern of each spot weld can be related to the reference curve. The key to welding process stability judgment lies in the construction of the reference curve and the selection of the similarity metric. In this work, the reference curve of each welding schedule is constructed by the low-rank and sparse decomposition method, and the similarity between measured signals and the reference curve is calculated by Euclidean distance.

3.1.1. Low-rank and sparse decomposition

Spot welds on the automotive production line are mostly qualified, so a series of DR signals of normal spot welds can be obtained through cluster analysis. A DR signal of normal spot weld can be assumed as $x = \text{ref} + \text{noise}$, ref is the reference curve of a welding schedule. Therefore, a data matrix composed of normal signals with the same schedule contains

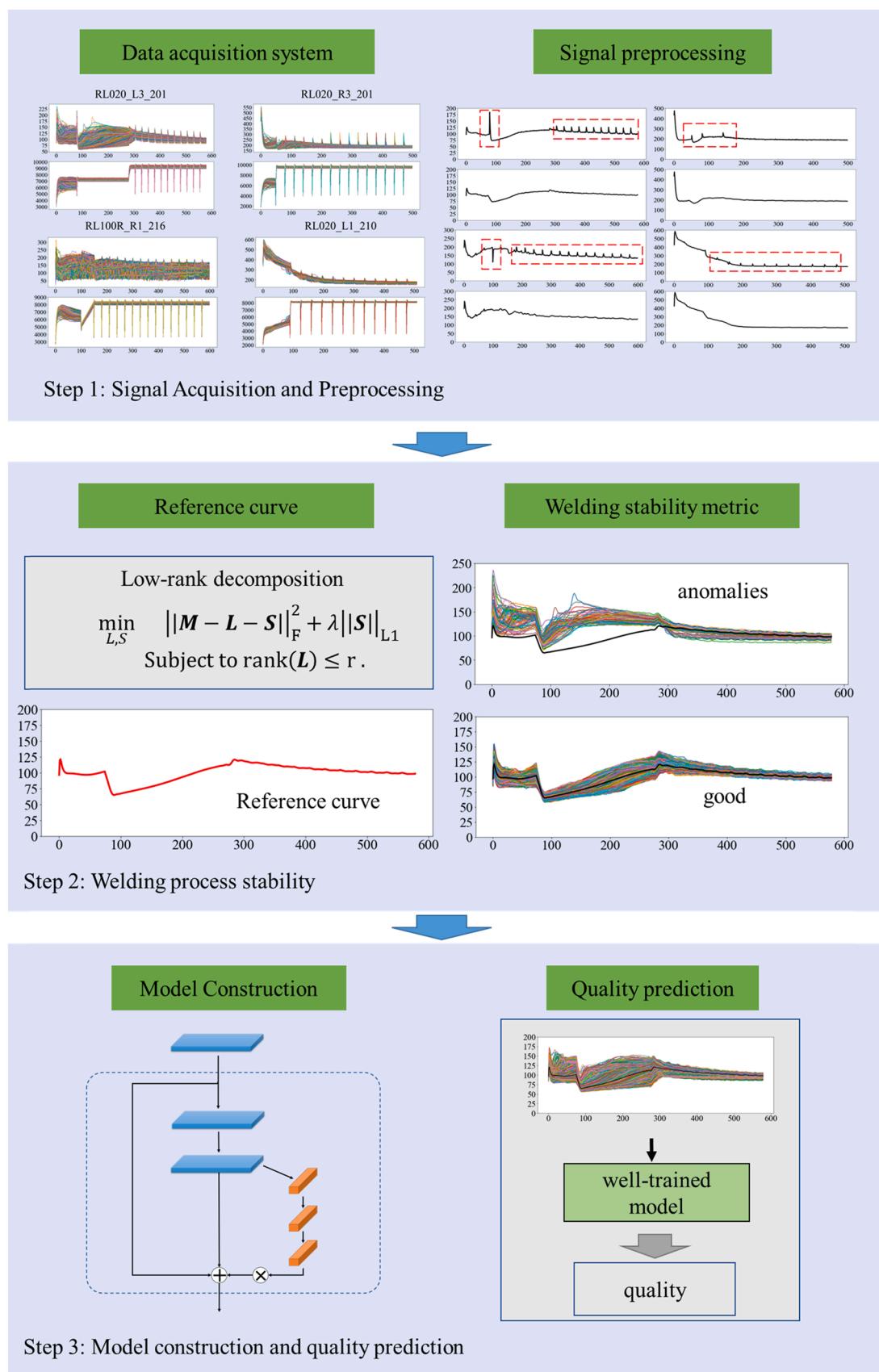


Fig. 5. The procedure for the proposed approach for spot welds quality evaluation.

a uniform variation pattern. The uniform variation pattern makes the values of each row of the data matrix close, resulting in a semi-low-rank matrix. The low-rank prior can be used to obtain physically interpretable information from the measured matrix. In the decomposed low-rank matrix, the *noise* caused by random disturbances has been removed, so that it can be better used to construct the reference curve of each welding schedule.

Mathematically, the measured signals can be organized as a data matrix $X \in \mathbb{R}^{M \times N}$ containing N -length signals recorded from M normal spot welds. From the above statement, we know that X can be expressed as $X = L + S$, where S is sparse noise, and L represents the low-rank matrix composed of the reference curves. Therefore, the goal is to extract low-rank components of the data matrix, which can be achieved by the low-rank and sparse decomposition method. Low-rank sparse decomposition of matrices is primarily implemented by robust principal component analysis (RPCA) [46] or Semi-Soft GoDec (SSGD) [47]. Due to the advantages of speed and capability of rank parameter selection, the SSGD algorithm is chosen to separate low-rank information from the signal matrix in this work. The SSGD algorithm can separate matrix X into three components defined in Eq. (1) using the minimization problem given in Eq. (2):

$$X = L + S + R, \text{rank}(L) \leq r, \quad \text{card}(S) \leq k, \quad (1)$$

$$\min_{L,S} \|X - L - S\|_F^2 + \lambda \|S\|_{L1} \text{s.t.} \text{rank}(L) \leq r \quad (2)$$

eral random projection, instead of the singular value decomposition method [48], which significantly reduces the time cost. In the algorithm, to estimate the low-rank component L_t of X at iteration t , from two random projections $Y_1^{M \times P} = X A_1$ and $Y_2^{N \times P} = X^T A_2$, a fast rank r approximation of X [49] has been used by independent Gaussian random matrices as observed in Eq. (3):

$$L_t = Y_1 (A_2^T Y_1)^{-1} Y_2^T \quad (3)$$

where $A_1 \in \mathbb{R}^{N \times P}$ and $A_2 \in \mathbb{R}^{M \times P}$ are random matrices. The sparse component (S_t) of X at iteration t can be estimated by applying a soft-thresholding operator with a threshold value equal to λ as expressed in Eq. (4):

$$S_t = \text{Soft_threshold}(X - L_t, \lambda) \quad (4)$$

The soft thresholding operator for matrix Q is defined as:

$$\text{Soft_threshold}(Q, \lambda) = \max(|q| - \lambda, 0) \text{sgn}(q) \quad (5)$$

where q denotes each element of the matrix Q and $\text{sgn}()$ is the sign function.

To sum up, the flowchart of SSGD is summarized in Algorithm 1.

Algorithm 1. : SSGD for signal matrix decomposition.

Input: signal matrix X and hyperparameter $r, \lambda, p, \varepsilon$

Output: low-rank matrix L and sparse matrix S

1. Initialize: $L_0 = X, S_0 = \mathbf{0}, t = 0$
2. while $\frac{\|X - L_t - S_t\|_F}{\|X\|_F} > \varepsilon$ do
3. $t = t + 1$
4. $\tilde{L} = [(X - S_{t-1})(X - S_{t-1})^T]^p (X - S_{t-1})$
5. $Y_1 = \tilde{L} A_1, A_2 = Y_1$
6. $Y_2 = \tilde{L}^T A_1 = Q_2 R_2, Y_1 = \tilde{L}^T A_2 = Q_1 R_1$
7. if $\text{rank}(A_2^T Y_1) < r$ then $r = \text{rank}(A_2^T Y_1)$, go to step 3; end;
8. $L_t = Q_1 [R_1 (A_2^T Y_1)^{-1} R_2^T]^{1/(2p+1)} Q_2^T$;
9. $S_t = \text{Soft_threshold}(X - L_t, \lambda);$
10. end while

where r and k are the maximum ranks of $L \in \mathbb{R}^{M \times N}$ and the maximum number of elements of $S \in \mathbb{R}^{M \times N}$, respectively. $\|\cdot\|_{L1}$ denotes $L1$ norm, defined as $\|S\|_{L1} = \sum_{ij} |s(i,j)|$, and $\|\cdot\|_F$ denotes Frobenius norm, defined as $\|P\|_F = \sum_{i=1}^m \sum_{j=1}^n |p(i,j)|^2$.

$R \in \mathbb{R}^{M \times N}$ is the approximation error and the hyperparameter λ is the trade-off between the error term and the sparsity of S .

The SSGD algorithm uses a low-rank approximation based on bilat-

3.1.2. Process stability metric

In order to measure the process stability of each spot weld, reasonable evaluation criteria need to be established. There are various ways to measure the similarity between two time series, such as Euclidean distance and dynamic time warping (DTW) distance. DR signals with the same welding schedule have the same length and the Euclidean distance is simple and fast to calculate. Therefore, based on Euclidean distance, the stability factor (SF) is defined as the deviation of the measured DR

signal from the corresponding reference curve by using Eq. (6).

$$SF = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - r_i)^2} \quad (6)$$

where x and r denote the measured signal and the corresponding reference curve, n is the length of DR signals.

During mass production, SF can be calculated in real-time and used for online quantitative evaluation of weld stability. The smaller SF , the closer the actual welding process to the reference curve, and the more stable the welding process. In this work, spot welds that are highly consistent with the reference curve are considered as good quality. Conversely, it means that the unstable situation occurred during the actual welding, and there is a risk of quality problems. It should be noted that the stability of a single welding process is not necessarily related to the welding quality, but the stability of the welding process for a large number of spots reflects the consistency of quality. By setting the threshold value of SF , the stable and unstable spots can be distinguished and alerted. Therefore, by counting the percentage of suspicious spots at the same position on different vehicles, we can identify the locations of spots with poor overall stability, thus guiding the production site to improve the welding process.

The welding process stability judgment has a clear physical meaning and is suitable for welding lines with the characteristics of mass production. By comparing the evaluation index of each spot weld with the qualified threshold value, the body spot welding process can be monitored online. At the same time, the evaluation results are presented in the form of the proportion of suspicious spot welds, which can directly represent the consistency of quality, and help identify the unstable factors on-site.

3.2. Deep learning-based quality prediction

As mentioned before, the results of the online stability judgment can be used as an a priori condition for other quality evaluation methods. It can help identify good spot welds that are highly consistent with the

reference curve and abnormal spot welds. In this section, deep learning models are introduced to predict the quality of spot welds that are difficult to judge by process stability.

During the welding process, pre-heating is used to melt the coating, while the main welding pulse forms the weld nugget. Nugget size and quality information are embedded in the time series between pulses. Therefore, the collected DR signals can be segmented into different channels for analysis according to the pulse time. The impact of each channel on the final quality prediction is different. For example, if there is a significant drop in one channel, then the spot weld is likely to be a burn-through. The quality prediction will be more accurate if the deep learning model can pay more attention to this channel. Therefore, the channel attention mechanism is naturally introduced to improve the prediction accuracy of the deep network model by extracting useful feature maps from different channels and automatically learning the weight of feature maps to enhance the expression of defect information.

The proposed network architecture is shown in Fig. 6. The network shares some of the same basic components as the traditional CNN models, including the convolutional (Conv) layer, rectifier linear unit (ReLU) activation function [50], batch normalization (BN) [51], and global average pooling (GAP) [52]. The Conv layer performs convolution operations with shared weights on the input signals to extract features and generate several feature maps. BN layer can accelerate the learning of the neural network and provide a weak regularization to alleviate overfitting. Due to sparsity, ReLU activation functions have less temporal and spatial complexity and avoid the vanishing gradient problem. GAP calculates the average value of each feature map separately to compress the global information into the channel description. Also, GAP can greatly reduce the weight parameters to be used by a fully connected (FC) layer to mitigate the occurrence of overfitting problems. Moreover, the essential component of the proposed network is the squeeze and excitation (SE) module. The SE module is a channel attention block that explicitly models channel-interdependencies within modules and selectively enhances important feature maps and suppresses less useful ones. Therefore, the weight of each feature map is automatically learned at each layer of the network and the benefits of

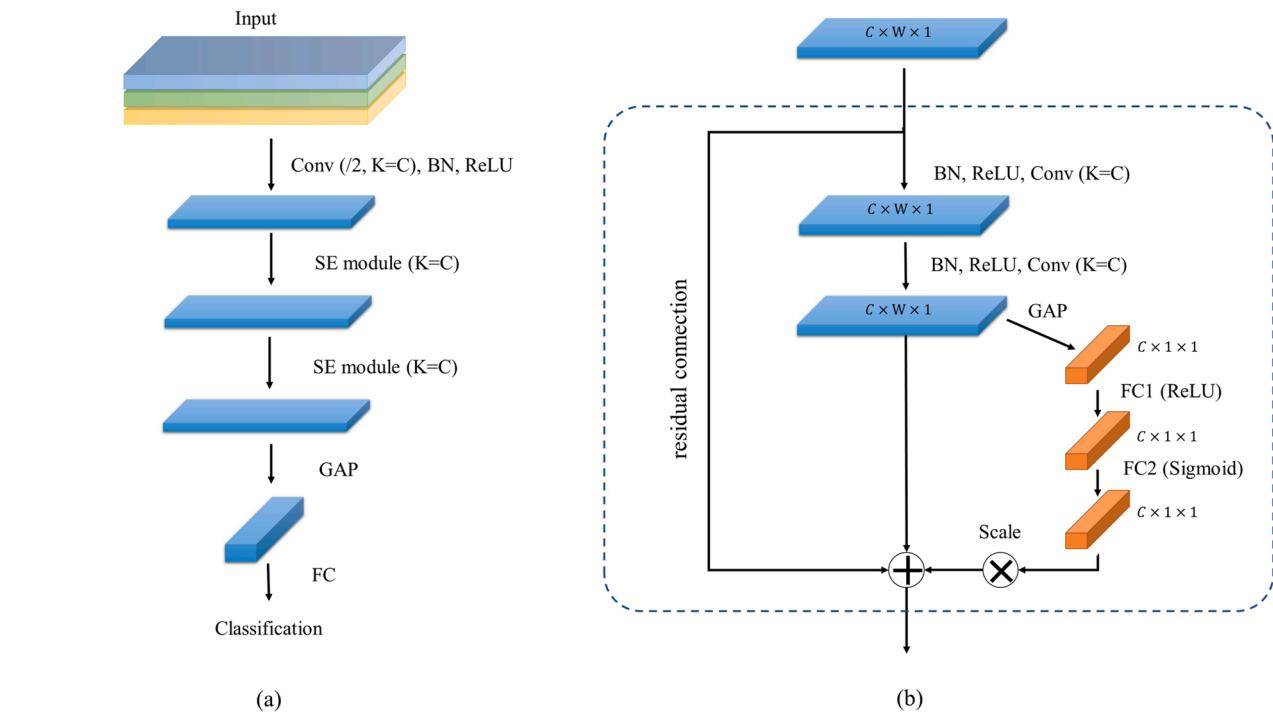


Fig. 6. (a) The proposed network architecture for spot welds quality evaluation. (b) SE module, where K is the number of convolutional kernels in the Conv layer, and C , W , and 1 in $C \times W \times 1$ represent the number of channels, width, and height of the feature map, respectively.

the SE block can be accumulated through the entire network to boost the feature discriminability.

In SE modules, the input is initially fed into convolutional layers to obtain the feature maps $X = [x_1, x_2, \dots, x_c]$, where the size of the feature map x_i is $W \times 1$, and the number of channels is C . Global information is squeezed to generate channel-wise statistics by using GAP operation:

$$z_c = \frac{1}{1 \times W} \sum_{i=1}^1 \sum_{j=1}^W X_c(i, j) \quad (7)$$

The squeezed feature map z_c represents the compressed value of the c -th channel of the input feature map, and its shape is $C \times 1 \times 1$. Then, two FC layers are utilized to learn the weight of each channel. The activation function of the first FC layer is ReLU, which can be mathematically expressed as:

$$S_{FC1} = \max(W_1 \times z + b_1, 0) \quad (8)$$

where W_1 and b_1 denote the weight and bias of the FC layer, respectively. The second FC layer maps the value to 0–1 to calculate the weight of each channel through the sigmoid function:

$$S = \frac{1}{1 + e^{-(W_2 \times s_1 + b_2)}} \quad (10)$$

where W_2 and b_2 denote the weight and bias of the FC layer, respectively.

Finally, the obtained weights are multiplied with the original feature maps X to achieve channel-wise feature recalibration. The channel-wise multiplication can be formulated as:

$$\tilde{x}_c = [s_1 \times x_1, s_2 \times x_2, \dots, s_i \times x_i, \dots, s_c \times x_c] \quad (10)$$

In addition, the residual connection structure is proved to have better performance than the traditional CNN structure [53]. The residual structure can unleash the feature learning capability of deep neural networks. Thus the output of the channel attention mechanism combined with the residual connection has the form:

$$H(x) = \tilde{x}_c + x \quad (11)$$

The spot welds quality prediction can be expressed as a classification problem. Deep learning models learn the network parameters by minimizing the loss function, thereby establishing a complex mapping relationship. Among them, cross-entropy error is the most commonly used loss function in classification tasks [50]. Providing the labeled data, the model can learn the significant features that represented different classes. Because the DR signals datasets collected on the actual production lines is imbalanced, the weighted cross-entropy loss function is selected for the quality prediction problem, and it can be formulated as:

$$L = - \sum_{j=1}^{N_{class}} w_j t_j \log(y_j) \quad (12)$$

where t_j and y_j denote the actual and predicted probability of j -th class, respectively. w_j is the weight coefficient of each class. After calculating the weighted cross-entropy error, the gradient descent algorithm can be used to optimize the network parameters, and the well-trained model can be used for online quality inspection.

4. Results and discussion

The experimental studies including welding process stability and deep learning-based classifiers are carried out to demonstrate the performance of the proposed quality inspection approach in this section. To analyze the generalization and universality of the approach, several datasets collected on production lines with different welding schedules are selected.

4.1. Reference curve construction

The key to welding process stability judgment lies in the construction of the reference curve and the selection of the similarity metric. The measured signal matrix is composed of a series of normal signals with the same welding schedule. [Algorithm 1](#) is employed to extract the low-rank components of the measured signal matrix, which can be used to construct the reference curve. Here, the hyperparameters in the [Algorithm 1](#) are set to $p = 6$, $r = 1$, and $\lambda = 0.5$. [Fig. 7](#) presents the eigenvalue distribution of the low-rank and sparse matrix obtained by [Algorithm 1](#). It can be found that the decomposed low-rank matrix well meets the low-rank characteristics because the algorithm imposes a hard constraint on the rank. Illustrations of each DR signal decomposition and the corresponding reference curve under different welding schedules are shown in [Figs. 8–11](#). It can be drawn that the extracted low-rank components are very close to the corresponding uniform variation pattern, indicating that physically meaningful components from the measured signal matrix are indeed obtained. The low-rank component captures the general trend as reflected in DR signals of good quality spot welds, which can be used as a reference curve under the welding schedule.

The obtained reference curve of each spot weld is used to conduct a preliminary screening and evaluation of the collected signals. By setting the threshold value, spot welds whose DR curve obviously deviates from the reference curve can be screened out, which is regarded as abnormal and output alarm. In addition, spot welds that are highly consistent with the reference curve are considered to be of good quality in this quality inspection framework. The filtered good and abnormal signals under different schedules are shown in [Fig. 12](#). Abnormal signals can be easily eliminated by the corresponding reference curve so that it will not enter the next step of data-driven quality judgment and affect its training process.

4.2. Performance assessment of deep learning models

4.2.1. Dataset and evaluation criterions

Nugget size and welding quality information are contained in the time series between pulses. Therefore, the quality of spot welds can be predicted by analyzing the information within each pulse. DR signals were segmented into discrete windows according to pulse time. [Fig. 13](#) shows a diagram of window segmented-based subsequence extraction. The window size is equal to the pulse duration of the weld schedule. With the window segmented technique, collected DR signals can be segmented into several channels.

Four spot welding datasets with different schedules on automotive production lines are selected to evaluate model performance. Some example DR signals with different schedules in the datasets are shown in [Fig. 1](#). The training and testing datasets are divided according to 8:2, and the category ratios are guaranteed to be consistent, as presented in [Table 1](#). As illustrated in [Table 1](#), it is a class-imbalance dataset where most samples belong to normal spot welds, while “cold welds” account for 0.54% and “burn through” only accounts for 0.14% by contrast.

It has been already known that using a single classification metric for class-imbalanced datasets can be misunderstanding. In order to better prove the defect identification ability of the proposed method, common classification metrics like false discovery rate (FDR), Recall and Accuracy are used to evaluate the quality prediction performances. The definitions of all these metrics are elaborated below and the relevant components are explained in [Table 2](#).

Note: TP: The model correctly predicts positive class as a positive class; FN: The model incorrectly predicts positive class as a negative class; FP: The model incorrectly predicts negative class as a positive class; TN: The model correctly predicts negative class as a negative class.

$$FDR = \frac{FP}{FP + TP} \quad (13)$$

FDR measures the percentage of defective welds among predictions

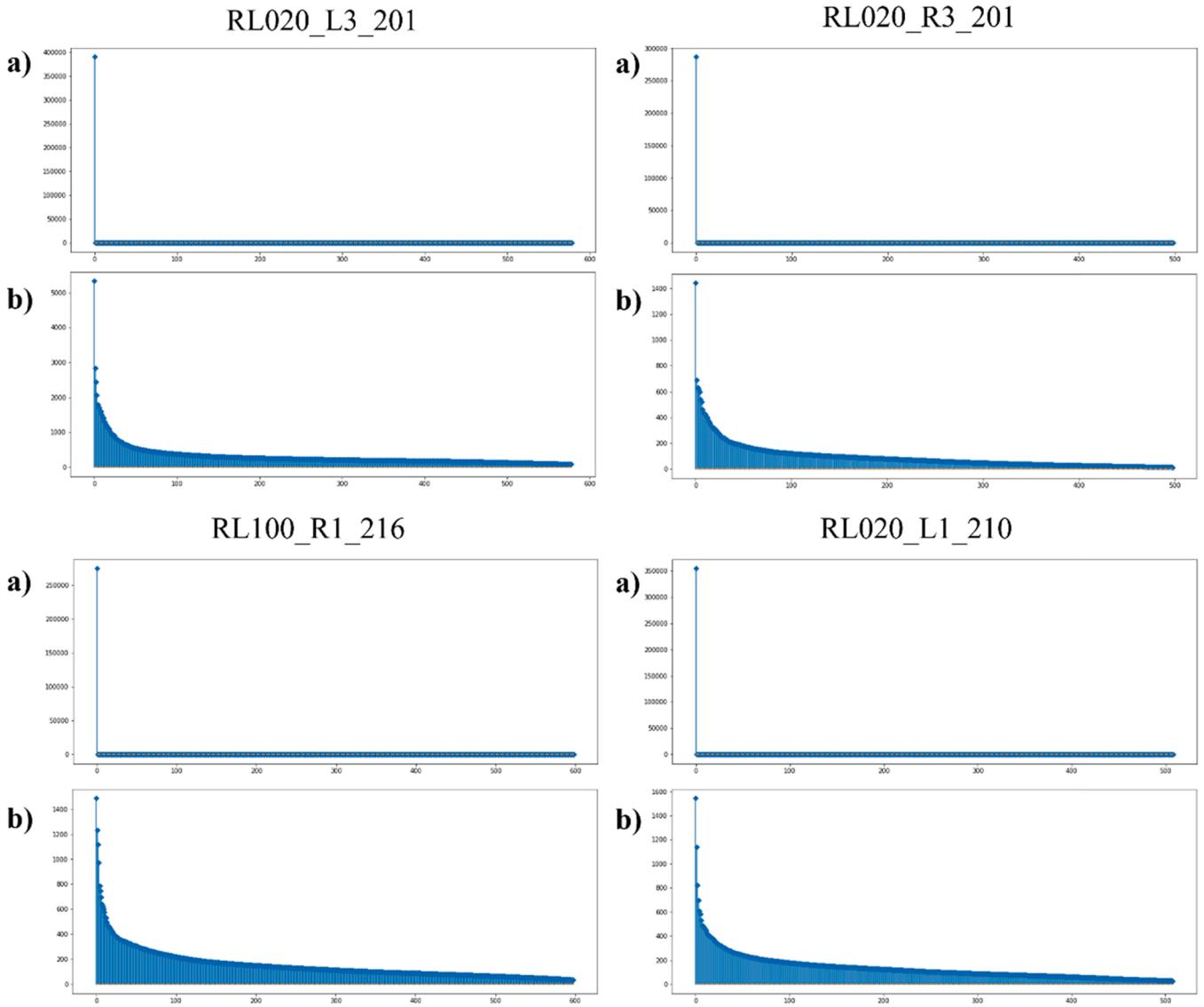


Fig. 7. The eigenvalue distribution of decomposed matrices. a) low-rank matrix; b) sparse matrix.

that the model considers to be of normal quality. This is a crucial indicator for automobile manufacturers because high FDR means that many defective spot welds will be considered normal, which will lead to catastrophic consequences. Besides, Recall indicates the possibility of truly normal spot welds identified by the model in all samples of real normal spot welds. Low Recall means that many normal spot welds are mistakenly detected as defects and singled out, resulting in wasted production costs and labor. Accuracy gives the overall performance of the model. Therefore, in the actual automotive production lines, the ideal situation is that the established quality prediction model has low FDR and high Recall.

4.2.2. Model comparison and analysis

For the present performance comparative study, several deep learning models such as ConvNet, ResNet [53], and LSTM [54] are selected to conduct the classifier. Since there is no consensus on how to set the hyperparameters of the deep network, this paper will set them based on several successful cases [41,42]. Meanwhile, the best hyperparameter is selected under a scheme of the five-fold cross-validation. The detailed architecture of the proposed network is as follows. The collected DR signal is segmented into several channels according to the pulse time and fed to the Conv layer, whose kernel number, size, and

stride of the convolution kernel are 16, 5, 2, respectively. Then, two SE modules are stacked after the Conv layer. As shown in Fig. 6, the kernel number, size, and stride of the Conv layer in the two SE modules are 16, 3, and 1 respectively. The output feature maps are reduced to 1D vectors after the GAP layer. In the end, the FC output layer has two or three neurons, which equals the number of considered classes. Moreover, the detailed training parameters of the experiments are given as follows. Adam [55] is implemented as the optimizer and the initial learning rate is 0.001. The number of training epochs is 100 and the mini-batch is set to 32. Finally, the best model evaluated by using the validation set during the optimization process is saved. The experiments are conducted on a computer with NVIDIA GeForce RTX 3060 GPU and the implementation is based on Tensorflow deep learning framework.

In order to clearly illustrate the misclassification results of each welding quality, the resulted confusion matrices of four total datasets are shown in Fig. 14. The larger the diagonal element value, the better the classification performance. It can be seen from the confusion matrices that the proposed model has good defect detection performance and can accurately identify spot welding defects such as cold welds and burn through. However, the performance of the proposed network model on the RL100_R1_216 and RL020_L1_210 datasets is mediocre. After careful analysis and field investigation, the reason is

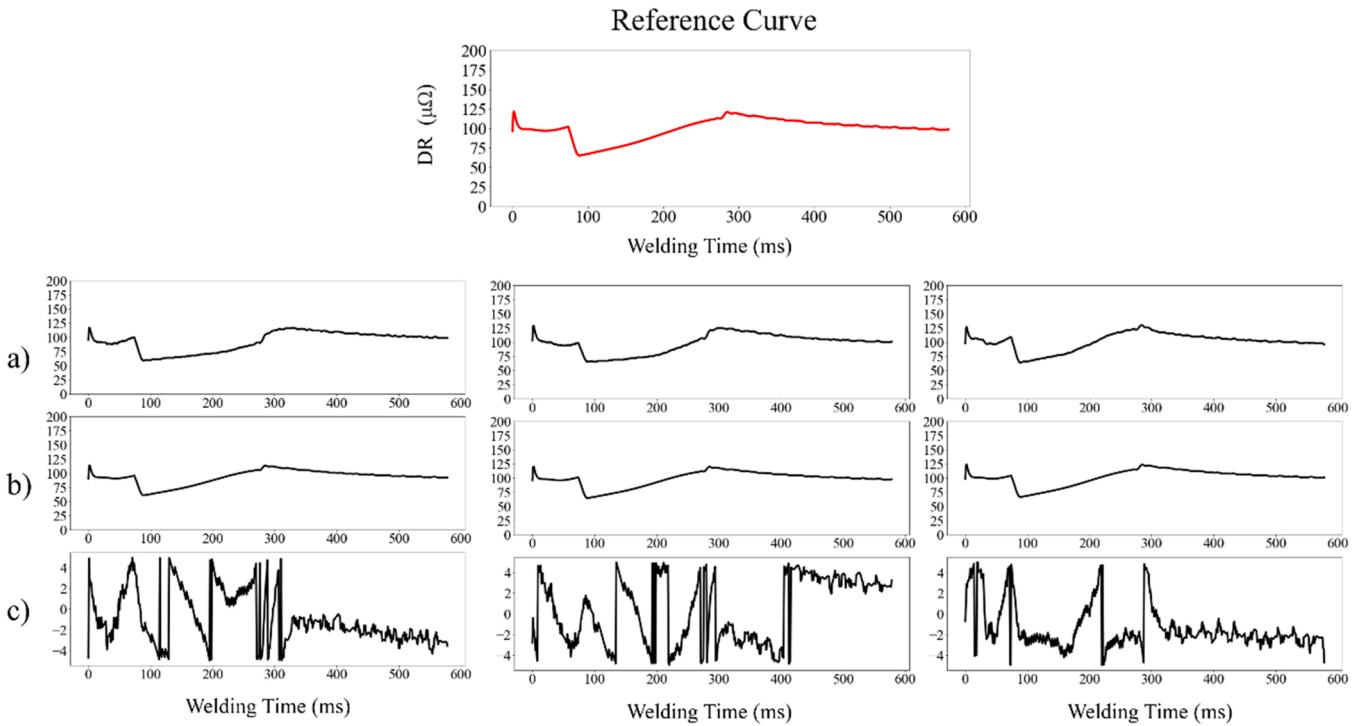


Fig. 8. Illustration of the decomposition of the signal obtained by using [Algorithm 1](#) and the corresponding reference curve in RL020_L3_201. a) measured DR signal; b) the extracted low-rank component; c) reflects noise component.

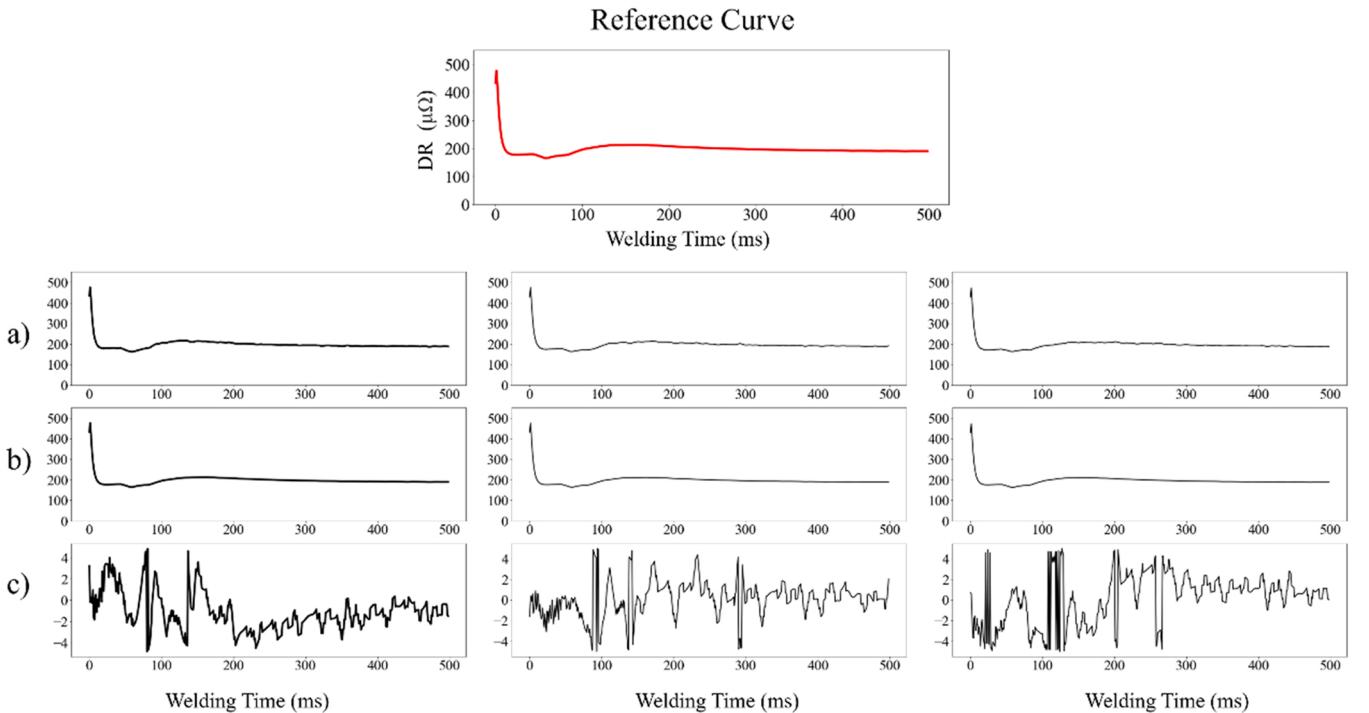


Fig. 9. Illustration of the decomposition of the signal obtained by using [Algorithm 1](#) and the corresponding reference curve in RL020_R3_201. a) measured DR signal; b) the extracted low-rank component; c) reflects noise component.

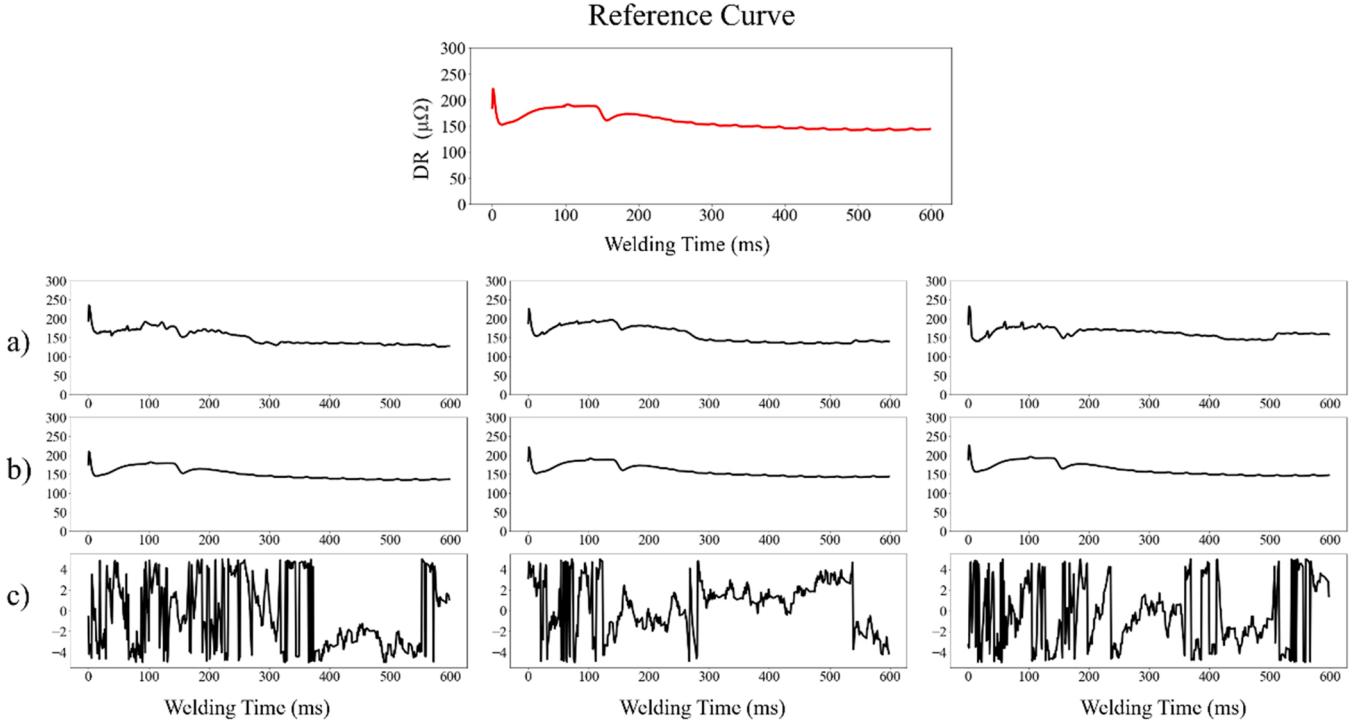


Fig. 10. Illustration of the decomposition of the signal obtained by using [Algorithm 1](#) and the corresponding reference curve in RL100_R1_216. a) measured DR signal; b) the extracted low-rank component; c) reflects noise component.

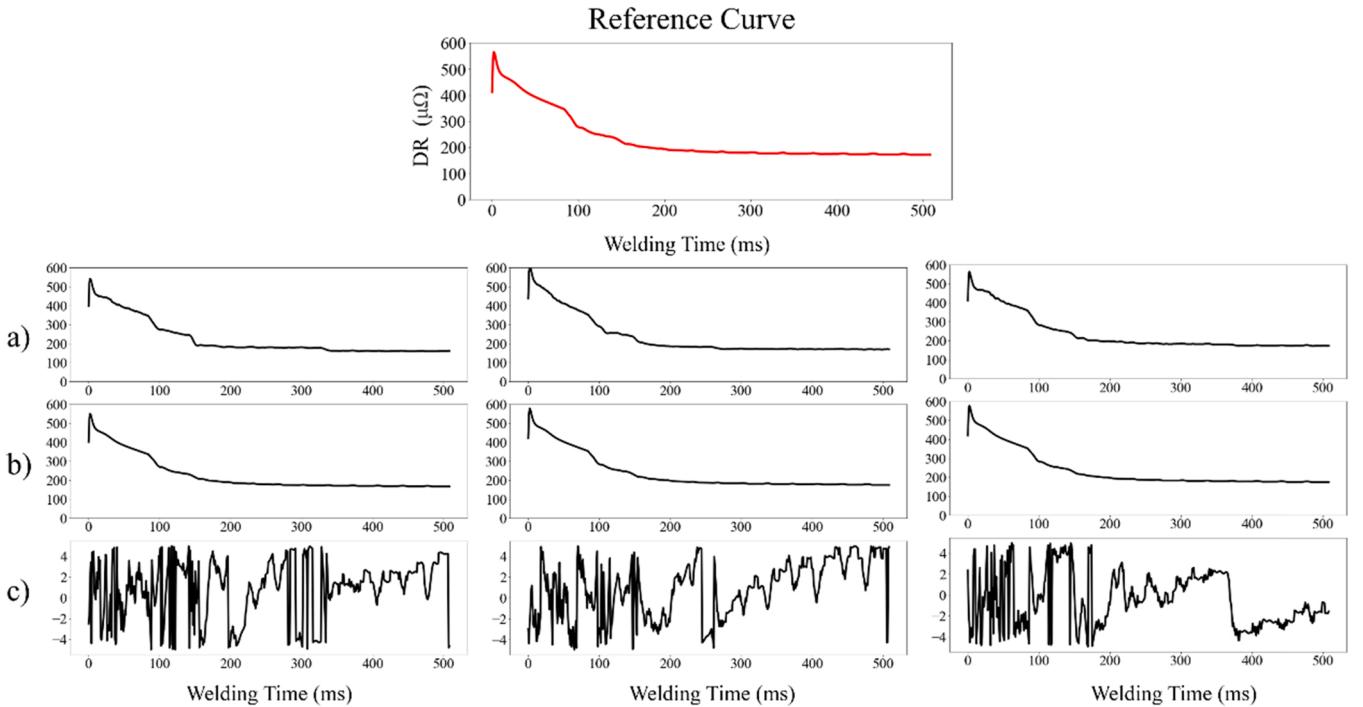


Fig. 11. Illustration of the decomposition of the signal obtained by using [Algorithm 1](#) and the corresponding reference curve in RL020_L1_210. a) measured DR signal; b) the extracted low-rank component; c) reflects noise component.

that the RL100_R1_216 spot welding process is not very stable in the field, resulting in the instability of the collected DR signal, as shown in [Fig. 1](#). The relatively poor data quality of RL100_R1_216 deteriorates the performance of the deep learning model. Meanwhile, the DR signal difference between small expulsion and burn through is slight, and both will have a sudden drop in amplitude of the curve. Small expulsion does not always destroy the strength of the weld, and in the actual production

lines, the small expulsion with an acceptable nugget size can be considered as a normal weld. As presented in the confusion matrix for RL100_R1_216 and RL020_L1_210 in [Fig. 14](#), some normal spot welds are wrongly identified as burn through by the trained model. A check of these signals revealed that most of them showed small expulsion.

Compared to misclassifying “burn through” as “cold weld”, it is more unacceptable to mistake “cold weld” for “normal”. As a result, all

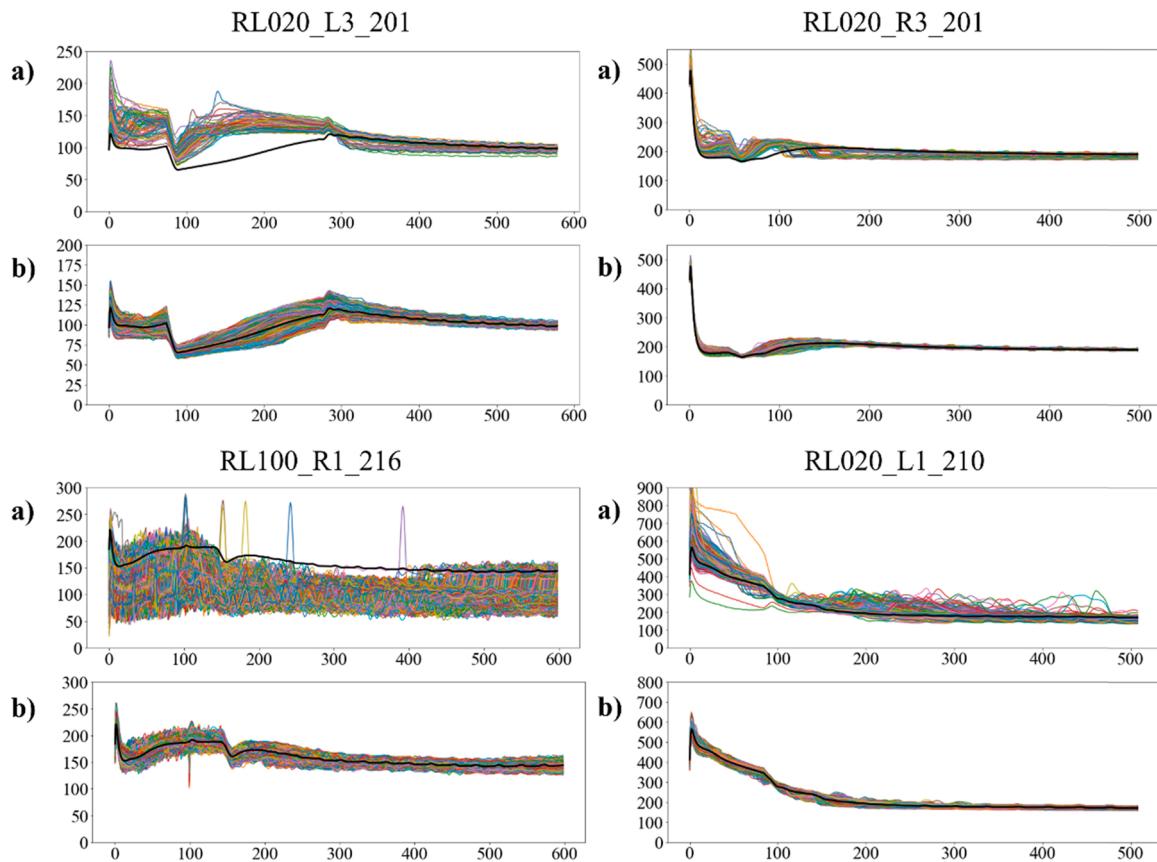


Fig. 12. Some filtered signals under different schedules. a) anomaly; b) good.

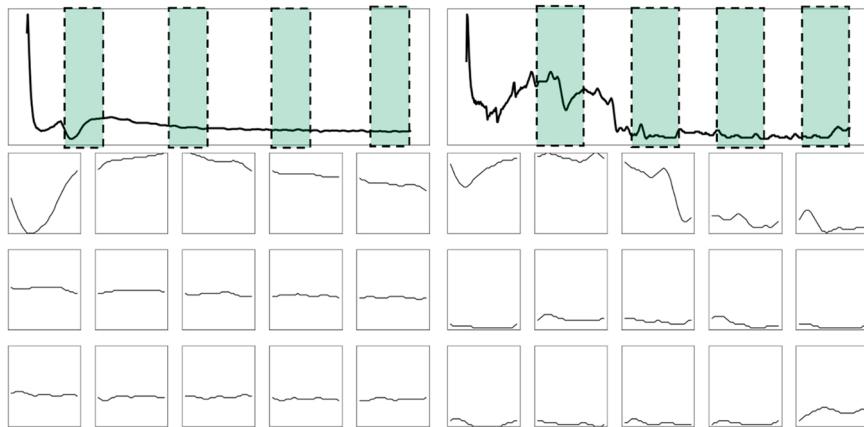


Fig. 13. Window segmented-based subsequence extraction.

Table 1
Number of training and testing samples in the datasets.

Schedules	Training set			Testing set		
	Normal	Cold weld	Burn through	Normal	Cold weld	Burn through
RL020_L3_201	11,318	116	0	2830	29	0
RL020_R3_201	6065	73	5	1517	18	1
RL100_R1_216	13,203	0	29	3302	0	7
RL020_L1_210	9305	27	24	2327	7	6

Table 2
Confusion matrix.

Confusion Matrix	Prediction	
	Positive	Negative
True	Positive Negative	True Positive (TP) False Positive (FP)
		False Negative (FN) True Negative (TN)

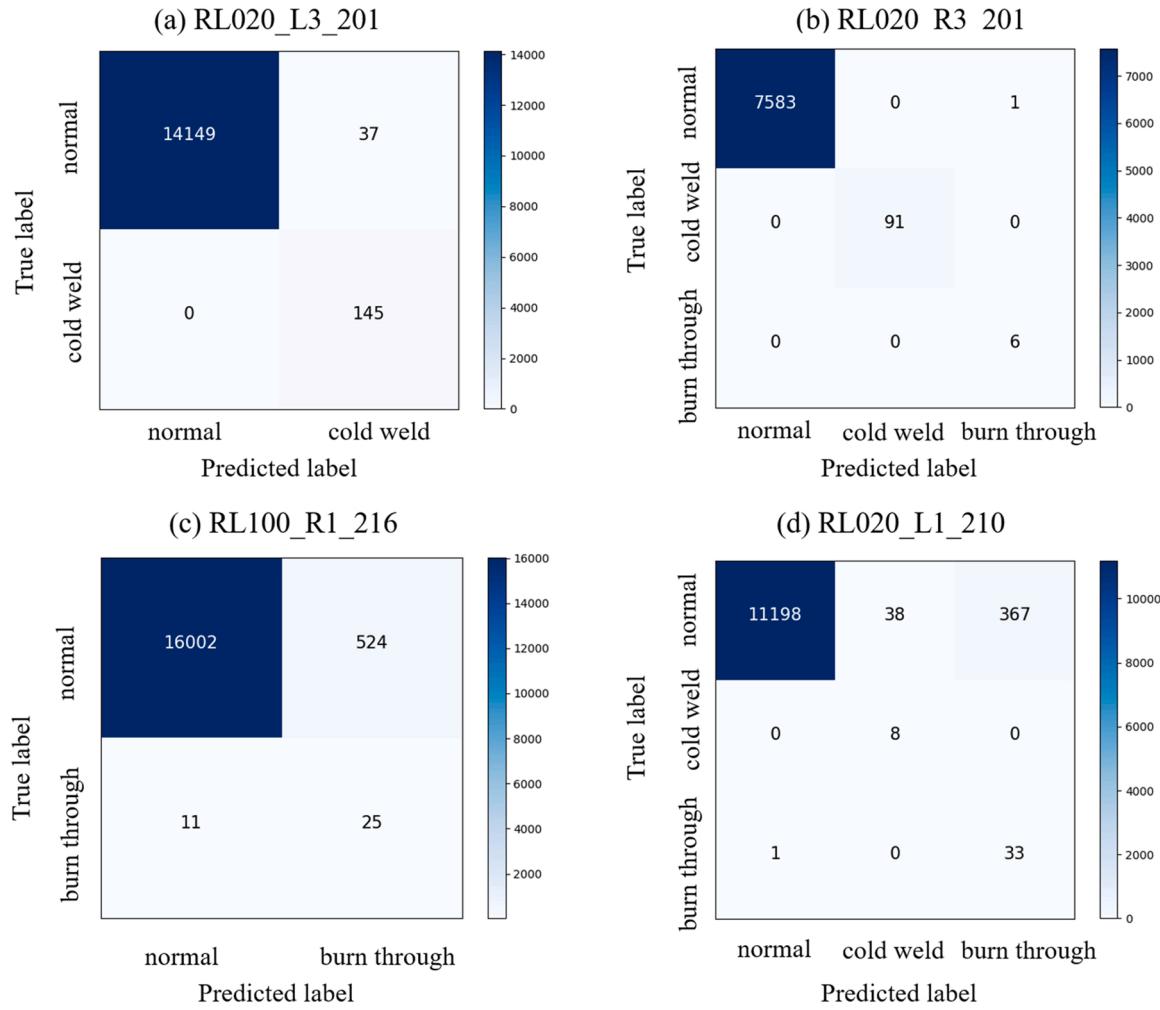


Fig. 14. Confusion matrices of the proposed model for different spot welding datasets.(a) RL020_L3_201; (b) RL020_R3_201; (c) RL100_R1_216; (d) RL020_L1_210.

Table 3

Comparisons of FDRs, Recall, Accuracy, and Time of different models on RL020_L3_201.

Model	FDR	Recall	Accuracy	Time (s)
ConvNet	0.11%	99.58%	99.47%	0.0177
ResNet	0.00%	98.52%	98.54%	0.0180
LSTM	0.00%	98.87%	98.88%	0.0174
Proposed model	0.00%	99.82%	99.82%	0.0181

Table 4

Comparisons of FDRs, Recall, Accuracy, and Time of different models on RL020_R3_201.

Model	FDR	Recall	Accuracy	Time (s)
ConvNet	0.00%	99.21%	99.22%	0.0134
ResNet	0.00%	99.87%	99.87%	0.0135
LSTM	0.00%	99.54%	99.54%	0.0140
Proposed model	0.00%	100.00%	100.00%	0.0138

defective spot welds are grouped as a “defective class”, which makes the calculation of FDRs and Recalls more straightforward. To provide more quantitative insights, Tables 3–6 illustrates comparison results of different deep learning models under different testing sets. The comprehensive comparison results demonstrate that the proposed model with the channel attention mechanism can reliably predict welding quality in different cases, and it makes considerable FDR

Table 5

Comparisons of FDRs, Recall, Accuracy, and Time of different models on RL100_R1_216.

Model	FDR	Recall	Accuracy	Time (s)
ConvNet	0.13%	91.50%	90.94%	0.0127
ResNet	0.19%	97.76%	97.59%	0.0131
LSTM	0.19%	94.95%	94.78%	0.0129
Proposed model	0.03%	95.88%	96.86%	0.0135

Table 6

Comparisons of FDRs, Recall, Accuracy, and Time of different models on RL020_L1_210.

Model	FDR	Recall	Accuracy	Time (s)
ConvNet	0.05%	87.51%	87.51%	0.0129
ResNet	0.05%	92.85%	92.83%	0.0131
LSTM	0.05%	92.33%	92.31%	0.0134
Proposed model	0.05%	96.47%	96.44%	0.0133

reduction and Recall increase compared with other models. The reason is that the integration of channel attention in the deep architectures can perform channel-wise feature recalibration, which enhances the expression of defect information and makes the learned high-level features more discriminative. The automatic learning mechanism of discriminative features improves the generalization and versatility of the quality evaluation approach. Furthermore, the quality prediction

times of the deep network models used in the experiment are compared, as presented in Tables 3–6. The runtime comparison is obtained by averaging all samples of the dataset. It can be concluded that the computational efficiency of the proposed network is sufficient for online quality inspection. The proposed approach has a good predictive ability for spot welding quality and provides an accurate online quality inspection solution for automotive production lines.

5. Conclusion

In the present work, an RSW quality inspection approach based on DR signals is proposed. To achieve a higher accuracy of quality prediction, welding process stability and a deep network with channel attention mechanism are developed. Low-rank and sparse decomposition is used to construct a reference curve and a similarity measure based on Euclidean distance is used to determine the process stability. Furthermore, the experimental results show that the proposed network makes considerable FDR reduction and Recall increase compared with several mainstream models on several RSW datasets collected on the actual production line. The study provides a new approach for spot welds quality prediction based on DR signals, which has some application prospects for automotive production lines.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to acknowledge the financial support of the National Natural Science Foundation of China (No. 52075325). The project is also supported by the "cross research fund for translational medicine" of Shanghai Jiaotong University (zh2018qnb17, zh2018qna37).

References

- [1] Xia YJ, Su Z, Lou M, Li Y, Carlson BE. Online precision measurement of weld indentation in resistance spot welding using servo gun. *IEEE Trans Instrum Meas* 2020;69(7):4465–75. <https://doi.org/10.1109/TIM.2019.2943981>.
- [2] Zhou K, Yao P. Overview of recent advances of process analysis and quality control in resistance spot welding. *Mech Syst Signal Process* 2019;124:170–98. <https://doi.org/10.1016/j.ymssp.2019.01.041>.
- [3] Podrážaj P, Jerman B, Simončič S. Poor fit-up condition in resistance spot welding. *J Mater Process Technol* 2016;230:21–5. <https://doi.org/10.1016/j.jmatprotec.2015.11.009>.
- [4] Xia YJ, Su ZW, Li YB, Zhou L, Shen Y. Online quantitative evaluation of expulsion in resistance spot welding. *J Manuf Process* 2019;46:34–43. <https://doi.org/10.1016/j.jmapro.2019.08.004>.
- [5] Pei FQ, Tong YF, Yuan MH, Ding K, Chen XH. The digital twin of the quality monitoring and control in the series solar cell production line. *J Manuf Syst* 2021; 59:127–37. <https://doi.org/10.1016/j.jmsy.2021.02.001>.
- [6] Wang Q, Jiao W, Zhang Y. Deep learning-empowered digital twin for visualized weld joint growth monitoring and penetration control. *J Manuf Syst* 2020;57: 429–39. <https://doi.org/10.1016/j.jmsy.2020.10.002>.
- [7] Wang Y, Perry M, Whitlock D, Sutherland JW. Detecting anomalies in time series data from a manufacturing system using recurrent neural networks. *J Manuf Syst* 2020. <https://doi.org/10.1016/j.jmsy.2020.12.007>.
- [8] Cai W, Wang JZ, Jiang P, Cao LC, Mi GY, Zhou Q. Application of sensing techniques and artificial intelligence-based methods to laser welding real-time monitoring: a critical review of recent literature. *J Manuf Syst* 2020;57:1–18. <https://doi.org/10.1016/j.jmsy.2020.07.021>.
- [9] Wang B, Hu SJ, Sun L, Freiheit T. Intelligent welding system technologies: state-of-the-art review and perspectives. *J Manuf Syst* 2020;56:373–91. <https://doi.org/10.1016/j.jmsy.2020.06.020>.
- [10] Dogan A, Birant D. Machine learning and data mining in manufacturing. *Expert Syst Appl* 2021;166:114060. <https://doi.org/10.1016/j.eswa.2020.114060>.
- [11] Bhole SD, Ma C, Khan MS, Chen D. A study of spot welding of advanced high strength steels for automotive applications. *J Iron Steel Res Int* 2011;18:724–9. <https://doi.org/10.1016/j.hydromet.2011.02.011>.
- [12] Zhou L, Zheng W, Li T, Zhang T, Zhu S. A material stack-up combination identification method for resistance spot welding based on dynamic resistance. *J Manuf Process* 2020;56:796–805. <https://doi.org/10.1016/j.jmapro.2020.04.051>.
- [13] Matsushita M, Ikeda R, Oi K. Development of a new program control setting of welding current and electrode force for single-side resistance spot welding. *Weld World* 2015;59:533–43. <https://doi.org/10.1007/s40194-015-0228-1>.
- [14] Zhao DW, Wang YX, Zhang L, Zhang P. Effects of electrode force on microstructure and mechanical behavior of the resistance spot welded dp600 joint. *Mater Des* 2013;50:72–7. <https://doi.org/10.1016/j.matdes.2013.02.016>.
- [15] Min J. Real time monitoring weld quality of resistance spot welding for the fabrication of sheet metal. *J Mater Process Technol* 2003;132:102–13. [https://doi.org/10.1016/S0924-0136\(02\)00409-0](https://doi.org/10.1016/S0924-0136(02)00409-0).
- [16] Xia YJ, Zhou L, Shen Y, Wegner DM, Carlson BE. Online measurement of weld penetration in robotic resistance spot welding using electrode displacement signals. *Measurement* 2020;168:108397. <https://doi.org/10.1016/j.measurement.2020.108397>.
- [17] Luo Y, Wan R, Xie X, Zhu Y. Expulsion analysis of resistance spot welding on zinc coated steel by detection of structure-borne acoustic emission signals. *Int J Adv Manuf Technol* 2016;84:1995–2002. <https://doi.org/10.1007/s00170-015-7846-z>.
- [18] Dejan A, Kurtov O, Rymanen PV. Acoustic emission as a tool for prediction of nugget diameter in resistance spot welding. *J Manuf Process* 2020;62:7–17. <https://doi.org/10.1016/j.jmapro.2020.12.002>.
- [19] Xia YJ, Li Y, Lou M, Lei H. Recent advances and analysis of quality monitoring and control technologies for resistance spot welding. *Chin J Mech Eng* 2020;31:100–25. <https://doi.org/10.3969/j.issn.1004-132X.2020.01.011>.
- [20] Zhao D, Bezugans Y, Wang Y, Du W, Vdonin N. Research on the correlation between dynamic resistance and quality estimation of resistance spot welding. *Measurement* 2021;168:108299. <https://doi.org/10.1016/j.measurement.2020.108299>.
- [21] Wan X, Wang Y, Zhao D, Huang Y, Yin Z. Weld quality monitoring research in small scale resistance spot welding by dynamic resistance and neural network. *Measurement* 2017;99:120–7. <https://doi.org/10.1016/j.measurement.2016.12.010>.
- [22] Wang X, Wang Y, Zhao D, Huang Y. A comparison of two types of neural network for weld quality prediction in small scale resistance spot welding. *Mech Syst Signal Process* 2017;93:634–44. <https://doi.org/10.1016/j.ymssp.2017.01.028>.
- [23] Xing B, Xiao Y, Qin QH, Cui H. Quality assessment of resistance spot welding process based on dynamic resistance signal and random forest based. *Int J Adv Manuf Technol* 2018;94:327–39. <https://doi.org/10.1007/s00170-017-0889-6>.
- [24] Zhang H, Wang F, Tao X, Jian Z, Wang L, Gao W. A novel quality evaluation method for resistance spot welding based on the electrode displacement signal and the Chernoff faces technique. *Mech Syst Signal Process* 2015;62:431–43. <https://doi.org/10.1016/j.ymssp.2015.03.007>.
- [25] Zhang H, Hou Y, Zhao J, Wang L, Xi T, Li Y. Automatic welding quality classification for the spot welding based on the hopfield associative memory neural network and Chernoff face description of the electrode displacement signal features. *Mech Syst Signal Process* 2017;85:1035–43. <https://doi.org/10.1016/j.ymssp.2016.06.036>.
- [26] Chen G, Sheng B, Luo R, Jia P. A parallel strategy for predicting the quality of welded joints in automotive bodies based on machine learning. *J Manuf Syst* 2022; 62:636–49. <https://doi.org/10.1016/j.jmsy.2022.01.011>.
- [27] Fawaz HI, Forestier G, Weber J, Idoumghar L, Muller PA. Deep learning for time series classification: a review. *Data Min Knowl Discov* 2019;33:917–63. <https://doi.org/10.1007/s10618-019-00619-1>.
- [28] Rui Z, Yan R, Chen Z, Mao K, Gao RX. Deep learning and its applications to machine health monitoring. *Mech Syst Signal Process* 2019;115:213–37. <https://doi.org/10.1016/j.ymssp.2018.05.050>.
- [29] Kiranyaz S, Avci O, Abdeljaber O, Ince T, Inman DJ. 1D convolutional neural networks and applications: a survey. *Mech Syst Signal Process* 2021;151:107398. <https://doi.org/10.1016/j.ymssp.2020.107398>.
- [30] Liu C, Hsiao W, Tu Y. Time series classification with multivariate convolutional neural network. *IEEE Trans Ind Electron* 2019;66(6):4788–97. <https://doi.org/10.1109/TIE.2018.2864702>.
- [31] Zhang Y, You D, Gao X, Zhang N, Gao PP. Welding defects detection based on deep learning with multiple optical sensors during disk laser welding of thick plates. *J Manuf Syst* 2019;51:87–94. <https://doi.org/10.1016/j.jmsy.2019.02.004>.
- [32] Zhang Z, Li B, Zhang W, Lu R, Wada S, Zhang Y. Real-time penetration state monitoring using convolutional neural network for laser welding of tailor rolled blanks. *J Manuf Syst* 2020;54:348–60. <https://doi.org/10.1016/j.jmsy.2020.01.006>.
- [33] Miao R, Shan Z, Zhou Q, Wu Y, Ge L, Zhang J, Hu H. Real-time defect identification of narrow overlap welds and application based on convolutional neural networks. *J Manuf Syst* 2021. <https://doi.org/10.1016/j.jmsy.2021.01.012>.
- [34] Liu T, Wang J, Huang X, Lu Y, Bao J. 3DSMDA-Net: an improved 3DCNN with separable structure and multi-dimensional attention for welding status recognition. *J Manuf Syst* 2021. <https://doi.org/10.1016/j.jmsy.2021.01.017>.
- [35] Sun L, Hu SJ, Freiheit T. Feature-based quality classification for ultrasonic welding of carbon fiber reinforced polymer through Bayesian regularized neural network. *J Manuf Syst* 2021;58:335–47. <https://doi.org/10.1016/j.jmsy.2020.12.016>.
- [36] Sun L, Tan C, Hu SJ, Dong P, Freiheit T. Quality detection and classification for ultrasonic welding of carbon fiber composites using time-series data and neural network methods. *J Manuf Syst* 2021;61:562–75. <https://doi.org/10.1016/j.jmsy.2021.10.010>.
- [37] Wang B, Li Y, Luo Y, Li X, Freiheit T. Early event detection in a deep-learning driven quality prediction model for ultrasonic welding. *J Manuf Syst* 2021;60: 325–36. <https://doi.org/10.1016/j.jmsy.2021.06.009>.

- [38] Ma D, Jiang P, Shu L, Geng S. Multi-sensing signals diagnosis and CNN-based detection of porosity defect during Al alloys laser welding. J Manuf Syst 2022;62: 334–46. <https://doi.org/10.1016/j.jmsy.2021.12.004>.
- [39] Candès EJ, Li XD, Ma Y, Wright J. Robust principal component analysis. J ACM 2011;58(3):1–37. <https://doi.org/10.1145/1970392.1970395>.
- [40] Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-excitation networks. IEEE Trans Pattern Anal 2017;42(8):2011–23. <https://doi.org/10.1109/TPAMI.2019.2913372>.
- [41] Xu X, Wang J, Zhong B, Ming W, Chen M. Deep learning-based tool wear prediction and its application for machining process using multi-scale feature fusion and channel attention mechanism. Measurement 2021;177:109254. <https://doi.org/10.1016/j.measurement.2021.109254>.
- [42] Zhao M, Zhong S, Fu X, Tang B, Pecht M. Deep residual shrinkage networks for fault diagnosis. IEEE Trans Ind Inf 2020;16(7):4681–90. <https://doi.org/10.1109/TII.2019.2943898>.
- [43] Wang B, Lou M, Shen Q, Li YB, Zhang HY. Shunting effect in resistance spot welding steels - Part1: experimental study. Weld J 2013;92(6):182s–189ss. <https://doi.org/10.1016/j.jmst.2013.03.006>.
- [44] Zhou L, Xia YJ, Shen Y, Haselhuhn AS, Wegner DM, Li YB, et al. Comparative study on resistance and displacement based adaptive output tracking control strategies for resistance spot welding. J Manuf Process 2021;63:98–108. <https://doi.org/10.1016/j.jmapro.2020.03.061>.
- [45] Hu S, Haselhuhn AS, Ma Y, Li YB, Carlson BE, Lin Z. Sensitivity of dissimilar aluminum to steel resistance spot welds to weld gun deflection. J Manuf Process 2021;68:534–45. <https://doi.org/10.1016/j.jmapro.2021.05.059>.
- [46] Tao M, Yuan X. Recovering low-rank and sparse components of matrices from incomplete and noisy observations. SIAM J Optim 2011;21:57–81. <https://doi.org/10.1137/100781894>.
- [47] Zhou T, Tao D. Godec: randomized low-rank & sparse matrix decomposition in noisy case. Proc ICML 2011:33–40.
- [48] Halko N, Martinsson PG, Tropp JA. Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. SIAM Rev 2011;53:217–88. <https://doi.org/10.1137/090771806>.
- [49] Fazel M, Candès E, Rechi B, Parrilo P. Compressed sensing and robust recovery of low rank matrices. IEEE 2008;10719045. ©10.1109/ACSSC.2008.5074571.
- [50] Ian G, Yoshua B, Aaron C. Deep learning. MIT Press; 2016. (<http://www.deeplearningbook.org/>).
- [51] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. Proc ICML 2015;37:448–56. (<https://arxiv.org/abs/1502.03167>).
- [52] Lin M, Chen Q, Yan S. Network in network. Proc ICLR 2014. (<https://arxiv.org/abs/1312.4400v1>).
- [53] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proc IEEE CVPR 2016:770–8. (<https://arxiv.org/abs/1512.03385>).
- [54] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8): 1735–80. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [55] Kingma DP, Ba JA. A method for stochastic optimization; 2014. (<https://arxiv.org/abs/1412.6980>).