



Comparison and explanation of data-driven modeling for weld quality prediction in resistance spot welding

Matthew Russell¹ · Joseph Kershaw² · Yujun Xia³ · Tianle Lv³ · Yongbing Li³ · Hassan Ghassemi-Armaki⁴ · Blair E. Carlson⁴ · Peng Wang^{1,2}

Received: 5 August 2022 / Accepted: 5 March 2023 / Published online: 1 April 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Resistance spot welding (RSW) is an important manufacturing process across major industries due to its high production speed and ease of automation. Though conceptually straightforward, the process combines complex electrical, thermal, fluidic, and mechanical phenomena to permanently assemble sheet metal components. These complex process dynamics make RSW prone to inconsistencies, even with modern automation techniques. This motivates online process monitoring and quality evaluation systems for quality assurance. This study investigates in-situ process sensing and neural networks-based modeling to understand key aspects of RSW process monitoring and offers three contributions: (1) a comparison of two data-driven modeling approaches, a feature-based Multilayer Perceptron (MLP) and a raw sensing-based convolutional neural network (CNN), (2) a comparison of how electrical and mechanical sensing data affect the model's performance, and (3) an explanation of MLP behavior using Shapley Additive Explanation (SHAP) values to interpret the contribution of sensing features to weld quality metric predictions. Both the MLP and CNN can predict weld quality metrics (e.g., nugget geometry) and detect a process defect (i.e., expulsion) using in-situ current and resistance sensing signals. Including force and displacement measurements improved performance, and the SHAP values revealed salient features underlying the RSW process (e.g., displacement contributes significantly to predicting axial nugget growth). Future work will explore additional architectural developments, explore ways to translate lab-developed models to production plants, and leverage these models to optimize RSW processes and improve quality consistency.

Keywords Resistance spot welding · Neural networks · Quality prediction · Process monitoring

Introduction

Resistance Spot Welding (RSW) is a vital joining process for sheet metal. From automotive bodies to aircraft frames, RSW is one of the oldest and most widespread automated manufacturing processes (Ma et al., 2013), clamping stacked metal

sheets between two electrodes and applying electrical current to heat, liquefy, and merge them into one permanent assembly. Instead of merging the entire mating surfaces into one contiguous component like a laminate, RSW creates localized spots of joined metal known as nuggets (Wang et al., 2020). These nuggets are analogous to rivets, screws, or other fasteners, but their mechanical and geometric properties lack the uniformity of prefabricated fasteners. Even though industrial welding robots can execute precise, repeatable operations, RSW is an inherently complex and inconsistent process. Minor changes in a material's chemical composition, structural and geometric properties (Ao et al., 2020), and sheet fit-up conditions, along with other process uncertainties, influence the weld quality by affecting the thermo-fluid dynamics. Other issues such as electrode tool degradation and unreliable control systems could further exacerbate process uncertainties. These process uncertainties lead to quality

✉ Peng Wang
edward.wang@uky.edu

¹ Department of Electrical and Computer Engineering, University of Kentucky, Lexington, KY, USA

² Department of Mechanical and Aerospace Engineering, University of Kentucky, Lexington, KY, USA

³ School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China

⁴ Global Research and Development, General Motors, Warren, MI, USA

inconsistencies and process defects that hinder quality assurance. This challenge motivates research into online quality evaluation using in-situ process sensing and predictive quality modeling to automatically gauge weld quality on the factory floor. Furthermore, modeling the latent connections between process parameters, sensing data, and the resulting weld quality is an essential step for future process optimization to improve quality consistency.

In open-loop RSW process control, quality assurance depends on this understanding of latent connections, or *process-quality relationships*, e.g., how material sheet fit-up conditions and process parameters affect weld quality. Some recent efforts attempted to model the complex RSW process dynamics through Finite Element Analysis (FEA) (Chen et al., 2018; Vignesh et al., 2019). However, the process-quality relationship modeling through FEA simulations is subject to many assumptions and neglects the necessary consideration of process uncertainties, making FEA inapplicable to real-world welding production lines. In-situ process monitoring provides a way to characterize process uncertainties and hence can supplement process parameters to improve process-quality relationship modeling and online prediction of weld quality (El-Sari et al., 2021). Since RSW consists of electrical, thermal, fluidic, and mechanical phenomena, in-situ monitoring can be approached by measuring different process aspects. Commonly investigated in-situ sensing includes welding current, Dynamic Resistance (DR), force, and electrode displacement. Temperature, geometry, and chemical properties change material resistance, so measuring DR captures important thermo-fluidic material fluctuations during joining (Wang & Wei, 2000). Force and displacement measurements capture similar information from a mechanical perspective since the thermal effects also physically strain the material (Batista et al., 2020).

Malformed nuggets can be revealed through identifying abnormal signals caused by errant fit-up conditions or process anomalies. For example, changes in electrode displacement could reflect an expulsion—the sudden ejection of molten metal out of the nugget—since material is lost from the welding area (Xia et al., 2020). Similarly, the DR signal is sensitive to abnormal sheet positioning, as improper sheet fit-up conditions reduce the conductivity between the electrodes (Zhou et al., 2021). To identify what specific information from the in-situ sensing can identify process anomalies like these, one important task is extracting features from raw time series signals. These features, typically derived from descriptive extrema values that delineate key process events, are often manually defined using empirical knowledge. For example, time domain features (e.g., starting values, peaks, slope rates, areas under the curves) of resistance and displacement signals vary with different sheet fit-up conditions that include the edge proximity of the welding point and large gaps between sheets (Zhou et al., 2021). Maximum

electrode displacement and how quickly it occurs can detect shunting, a defect caused when the welding current partially flows through a previous nugget (Xing et al., 2018). Frequency domain features, extracted through the Fourier transform or wavelet transformations, have also been investigated. Wu et al. (2018) used frequency analysis of force measurements to detect expelled welds, and Lee et al. (2020) derived frequency features from DR to detect electrode misalignment. Beyond manually defined, event-based features, many data-driven techniques such as Principal Component Analysis (PCA), isometric mapping, locally linear embedding (Zhao et al., 2020a), and Linear Discriminant Analysis (LDA) (Zhou et al., 2020) can compute abstract, dimensionless features that are useful for assessing process anomalies. Thus, with the proper features, predictive models can diagnose process inconsistencies and anomalies that affect RSW quality.

While basic process anomaly detection is valuable, predicting weld quality metrics from the in-situ sensing signals would be more beneficial. Xia et al. (2021) demonstrated that the difference between initial and peak displacement along with the displacement difference between the point of heating termination and the point of electrode removal could quantitatively predict the weld penetration through an empirical model. Another study analytically linked 20 features extracted from the DR signal to the weld nugget strength, and the mean DR value was found to be most reflective of nugget solidification (Zhao et al., 2021). These studies support the idea that a predictive model can estimate weld quality metrics from relevant sensing features. However, the explicit causal relationship between sensing features and weld quality is typically nonlinear and difficult to describe in terms of physical and empirical models. Thus, many data-driven predictive models leverage more advanced Machine Learning (ML) techniques. Zaharuddin et al. (2017) demonstrated an adaptive neuro-fuzzy interface system that predicted weld strength using manually-extracted DR features. Random forest was applied by Xing et al. (2017) to classify the DR signals to differentiate between cold, expelled, or good welds. Both time-domain and frequency-domain features were extracted from the DR curves and fed into a Support Vector Machine (SVM) and probabilistic neural network to predict weld quality by Lee et al. (2020). The results demonstrated that frequency-based features are more effective in detecting misalignment, while the highest prediction performance came from a combination of both feature domains. Combining DR features with welding parameters, El-Sari et al. (2021) developed a Multilayer Perceptron (MLP) model to predict nugget size, and data-driven features outperformed manually defined features. Similarly, Zhao et al. (2020b) showed that DR features extracted by PCA surpassed manually defined features as inputs for an extreme learning model predicting nugget size. These promising studies strongly suggest the

effectiveness of ML-based models for predicting RSW weld characteristics when provided with salient features.

Some ML models work directly with the raw time series sensing data, automatically determining relevant features. Wang et al. (2017) investigated recurrent neural networks, a Deep Learning (DL) method, to estimate the size of the Heat Affected Zone (HAZ) from raw DR and power curves. Also leveraging neural networks, Hwang et al. (2018) took current and voltage time series data as the inputs for an adaptive resonance theory-based model to predict strength, nugget size, and fracture shape. This evidence indicates that ML-based processing of raw sensing data could discover and retain critical information that improves quality prediction yet would be lost by manually defined features. The tradeoff to this method is interpretability. ML models, especially DL networks that could contain hundreds of thousands of parameters, abstract the raw sensing data into features that are no longer grounded in physical explanations. This can make them less informative for understanding RSW process dynamics. Thus, while automatic feature learning methods could achieve high accuracy, simpler models using physically-meaningful features could be more explainable.

Recognizing the ongoing need to study the capabilities of feature-based and raw sensing-based predictive models for RSW process monitoring and quality assurance, this study investigates the performance of neural network (NN) models for predicting weld quality and detecting expulsions under varying data conditions. The contributions of the paper's experiments and data analysis are threefold:

1. performance comparison of a feature-based Multilayer Perceptron (MLP) model and a raw sensing-based Convolutional Neural Network (CNN) when predicting key RSW quality metrics,
2. comparison of prediction performance between limited sensing (current and DR measurement, available in a production plant) and comprehensive sensing (i.e., current, DR, force, and displacement), and.
3. an interpretable explanation of the feature-based MLP model's data-driven predictions via each feature's contribution to output quality metrics, quantified by Shapley Additive Explanations (SHAP).

Comparing MLP and CNN models reveals the advantages and disadvantages of manual feature extraction and automated feature learning with deep NNs. The performance of these techniques can be domain dependent, so this study experimentally evaluates their suitability for RSW processes and provides guidance for future studies on in-situ process sensing for NN-based quality prediction and process anomaly detection. Furthermore, understanding how these models behave on subsets of the sensing variables can inform

smart factory installations and upgrades for cost effective solutions. Finally, the interpretability of NN models should be addressed to reconcile data-driven models and physical understanding, thereby increasing confidence in data-driven models that may be deployed in production environments. SHAP analysis and Shapley values provide this interpretability through a theoretically sound methodology that assigns model prediction contributions to the input features.

The remainder of this paper introduces the concepts of in-situ process monitoring and features in "[In-situ process monitoring in RSW](#)" section, surveys data-driven weld quality prediction and SHAP values in "[Data-driven weld quality metric prediction](#)" section, describes the welding experiments and NN training in "[Experiments](#)" section, presents results and discussions in "[Results and discussion](#)" section, and offers concluding remarks in "[Conclusion](#)" section.

In-situ process monitoring in RSW

In-situ sensing from the RSW process captures process dynamics and uncertainties. The following section details commonly studied sensing signals and explains their key characteristics with respect to the physical phenomena of RSW. Physically meaningful features are further chosen for training the MLP model, along with a discussion comparing feature-based and raw sensing-based analysis.

In-situ RSW sensing data

Commonly studied in-situ sensing for RSW process monitoring include current, DR, electrode force, and electrode displacement, since these electrical and mechanical signals could capture part(s) of the compound process dynamics. Figure 1 shows representative in-situ signal curves for DR, force, and displacement sensing while highlighting key points. Initially, DR decreases as the applied heat causes surface contaminants to breakdown, effectively removing resistance to the electric current (the initial point to A). The resistance then rises due to the increased temperature of the workpiece (point A to B) and slightly decreases when the sheets are mechanically and electrically joined together and a welding nugget forms (point B to C) (Wang & Wei, 2000). The DR signal drops out when the electrical current has been turned off to let the nugget solidify (point C to the end). Force and displacement curves have similar physically explainable points. During the initial stages of welding, the heat causes the workpiece to expand, increasing the electrode displacement and clamping force (point A to B). The metal then softens, and the force of the electrode compresses the workpieces. This plastic compression counteracts thermal expansion, leading to a largely stable signal (point B to C) (Ji & Zhou, 2004). Finally, both force and displacement

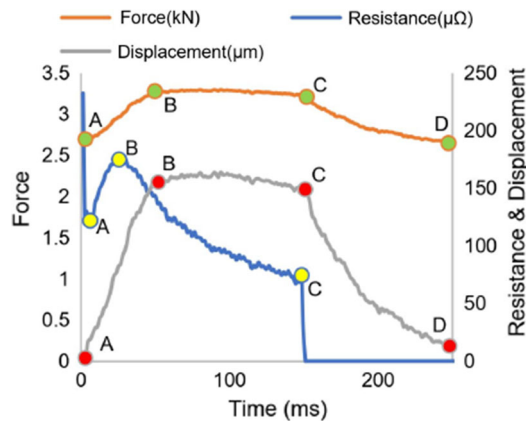


Fig. 1 Sample in-situ RSW process signals with key points labeled

trend downwards when heat is removed and the metal cools and contracts (point C to D) (Batista et al., 2020).

Physically meaningful RSW sensing features

While some studies include data-driven statistical features (e.g., PCA) extracted from the in-situ sensing data, physically meaningful features derived from the key points of the sensing data curves (see Fig. 1) facilitate building explainable models with clear connections to the underlying process. While statistical properties like variance, kurtosis, and higher-order moments could be used, these features best suit high-frequency periodic signals and lack the ability to capture the distinct sections and trends of welding curves, as shown in Fig. 1. To capture the welding process trends, physically explainable features extracted from the DR curve could be the average, final, and minimum resistance value during heating (point A to C) (Zhao et al., 2021). Features extracted from the displacement could include the change from the initial to the maximum value and the difference from the end of heating (point C) to solidification (point D) (Xia et al., 2021). Based on this preliminary work, this study derives a set of features from the key points of the curves to capture a holistic view of the welding process. For the DR signal, the extracted features include the initial value at point A, the difference and rate of change from point A to B, and the difference between points B and C. For force and displacement curves, the extracted features include the initial value at point A, the difference and rate of change from point A to B, the difference between point B and C, and the difference and rate of change from point C to D. The average measured current completes the set of 17 sensing features which is then combined with the two process parameters of welding current and time to total 19 physically meaningful features (see Table 1).

Table 1 Summary of input features

Signal	Features
Welding parameters	Welding time, welding current
Current	Mean
Dynamic resistance	A, ΔAB , ∇AB , ΔBC
Force	A, ΔAB , ∇AB , ΔBC , ΔCD , ∇CD
Displacement	A, ΔAB , ∇AB , ΔBC , ΔCD , ∇CD

Δ denotes the difference and ∇ denotes rate of change between the points

Feature-based versus raw sensing-based models

Both feature-based models and raw sensing-based models have their advantages and disadvantages. As explained previously, manually defined features convey key point information of electrical and mechanical signals and could align well with process dynamics. For example, the maximum delta force (i.e., the first-order signal difference) is a good indicator of expulsions, since the clamping force decreases sharply due to the sudden expulsion and loss of internal material, as shown in Fig. 2. The maximum delta force of all normal welds is below 0.15 kN, while that of expelled welds exceeds 0.2 kN. Therefore, this physically motivated feature design strategy can rely on more than abstract statistical correlations in the data. However, although manually defined features offer physical interpretability and generalizability, extracting them may cause information loss. In other words, the manually defined features cannot retain all the information contained in the raw data. Raw sensing-based models instead take the raw time series as the input and learn to extract features that are relevant to the prediction targets from a data-driven perspective. While these models tend to achieve better performance than feature-based models (Zhao et al., 2020b), the automatically extracted features lack physical interpretability. Hence, this study compares feature-based models and raw sensing-based models, the two major approaches for data-driven weld quality metric prediction.

Data-driven weld quality metric prediction

NN models, especially the DL variants, are state-of-the-art ML techniques. Consequently, this study investigates two NN models, MLP and CNN, to realize feature-based and raw sensing-based data-driven weld quality prediction and expulsion detection in RSW. Most NNs act as black-box models, but SHAP values can provide some insight into each feature's contribution to the quality prediction.

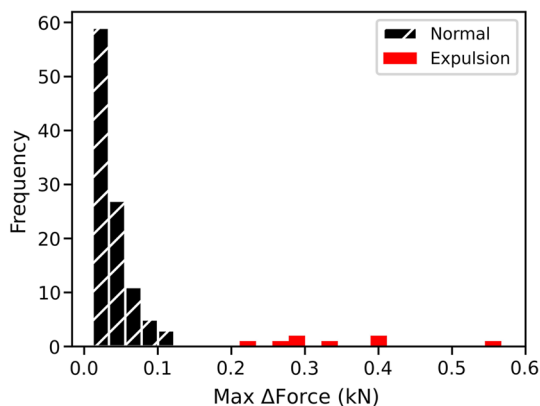


Fig. 2 Histogram of maximum change in welding force showing clear separation between welds without expulsions (< 0.15 kN) and those with expulsions (> 0.2 kN)

Multi-layer perceptron (MLP)

MLP represents the earliest and most widespread NN structure for linking discrete inputs to discrete outputs. MLP connects layers of artificial neurons with nonlinear activation functions to approximate complex input-output relationships. The input-output relationship in a three-layer perceptron can be expressed as $\mathbf{y} = f(W^{(2)}f(W^{(1)}\mathbf{x}))$, where \mathbf{x} is the column vector of input features, \mathbf{y} is the output, and $W^{(1)}$ and $W^{(2)}$ represent the matrix of weights from the input layer to the hidden layer and from the hidden layer to the output layer, respectively. Function f adds nonlinearity to the modeling, and popular nonlinear activation functions include the sigmoid function, hyperbolic tangent function, and the Rectified Linear Unit (ReLU). additional hidden layers can be inserted to increase the network’s capacity for modeling nonlinear systems.

Training an MLP is the process by which the network weights are tuned to achieve the desired input-output function that is empirically specified by a set of input-output pairs. A loss function measures the discrepancy between the model predictions and the expected outputs and allows training to be formulated as an optimization problem:

$$\hat{\theta} = \arg \min_{\theta} E_{(\mathbf{x}, \mathbf{y}) \sim \mathcal{D}} [\mathcal{L}(\mathbf{y}, g_{\theta}(\mathbf{x}))] \tag{1}$$

If $g_{\theta}(\mathbf{x})$ represents the outputs predicted by a network with parameters θ (e.g., the weight matrices $W^{(1)}$ and $W^{(2)}$), the goal is to find θ that minimizes the expected value of the loss function \mathcal{L} evaluated across all pairs (\mathbf{x}, \mathbf{y}) of input features \mathbf{x} and ground truth labels \mathbf{y} sampled from the training data distribution \mathcal{D} . In other words, solving the optimization problem will find the weights that minimize the error between the network outputs and the desired outputs across the entire training data set. For regression problems (e.g., predicting welding quality metrics), Mean Squared Error (MSE) suffices for

the loss function, while Binary Cross Entropy (BCE) is a frequent choice for binary classification tasks (e.g., detecting an expulsion defect). While analytically intractable, the optimization solution to (1) can be approximated by combining *loss backpropagation* that numerically computes the gradient of the loss with respect to network parameters from the model outputs back to the inputs and *gradient descent* that iteratively applies this gradient to update the parameters: $\theta'_i = \theta_i - \eta \partial \mathcal{L} / \partial \theta_i$, where θ_i is the i th weight, η is the learning rate hyperparameter that controls the speed of parameter adjustment, and θ'_i is the updated weight value. Common gradient descent algorithms include Stochastic Gradient Descent (SGD) and more advanced variants like Adagrad, RMSProp, or Adam.

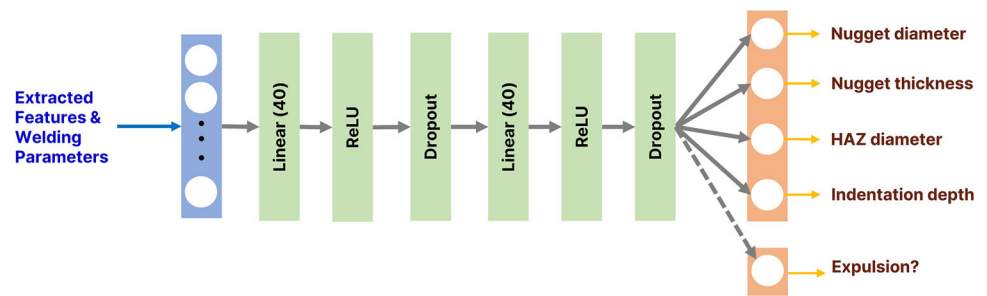
Figure 3 outlines an MLP structure with two hidden layers and ReLU nonlinear activation that either predicts RSW weld quality or detects expulsions from the 19 features specified in Table 1. While the architecture remains the same for both problems, a separate model is initialized and trained in each instance with either four regression outputs or one classification output. All four quality metrics are predicted using the same representation from the preceding layer, forcing this final hidden representation to summarize information about all four aspects of the weld while not imposing any explicit association or structure among the multivariate output. In both cases, the model takes normalized features as inputs to equally evaluate features with different ranges. Dropout layers mitigate overfitting by probabilistically masking layer outputs and forcing the network to rely on subsets of all the neurons.

Convolutional neural networks (CNN)

While MLPs can perform classification and regression tasks effectively, they remain dependent on the quality of the input features. Other NN and DL architectures, such as Convolutional Neural Networks (CNN), can automatically extract salient features from raw sensing data. CNN structures accommodate different data dimensionalities, and 1D-CNNs can process time series signals. A typical 1D-CNN structure consists of multiple 1D convolutional layers and 1D pooling layers for feature extraction and an MLP for mapping the extracted features to regression or classification outputs. 1D convolutional layers extract signatures and features by filtering time series signals with trainable 1D kernels. Intermediate outputs from successive layers represent more and more abstract features extracted from the input. A 1D convolutional operation can be expressed as

$$z_{i,j} = \sum_{u,v} w_{u,v}^{(j)} \cdot x_{si+u,v} \tag{2}$$

Fig. 3 The MLP structure used for weld quality prediction and expulsion detection, consisting of two hidden layers, ReLU activation, and dropout layers



where $w_{u,v}^{(j)}$ is the u th weight value for v th input channel in the j th kernel, $x_{si+u,v}$ is the $si+u$ th value of the v th input channel, and $z_{i,j}$ is the i th value of the j th output channel. Like MLP neurons, CNN neurons also apply a nonlinear activation function (usually ReLU) to z before passing it to downstream layers. As (2) illustrates, for $s > 1$ the CNN reduces the spatial dimension of the input, allowing later layers with the same kernel size to aggregate features into more abstract representations. Pooling operations, such as maximum pooling or averaging pooling, can further reduce the feature size:

$$z'_{i,j} = \max_{u=1,2,\dots,S-1} z_{si+u,j} \quad (3)$$

where S is the max pooling window size, s is the stride, $z_{si+u,j}$ is the $si+u$ th value of the j th channel, and $z'_{i,j}$ is the pooled value in the i th index of the j th output channel. Max pooling enables downstream layers to focus solely on the strongest patterns present in the input. After several successive 1D-CNN and pooling layers, the output is flattened into a feature vector \mathbf{z} which can be used by an MLP for the final classification or regression task.

Like training MLPs, training a CNN requires solving an optimization problem to find certain network parameters (including convolutional kernel weights and MLP weights) that minimize a task-dependent loss function (e.g., MSE or BCE). The gradient of the loss with respect to the convolution parameters is easily calculated, allowing CNNs to be trained with the same backpropagation and gradient descent algorithms used for MLPs. As a result, CNNs can be effective automatic feature extractors for detecting trace patterns within 1D time series data for a downstream classification and regression task.

Figure 4 shows a CNN structure that takes time series sensing data together with process parameters to predict weld quality metrics and detect expulsions. As with the MLP, each application initializes a separate instance of the same CNN architecture with either four regression outputs or one classification output, and no explicit structure is assumed among the four multivariate regression outputs besides using the same hidden representation from the previous layer. Three convolutional blocks consisting of 1D convolution with a

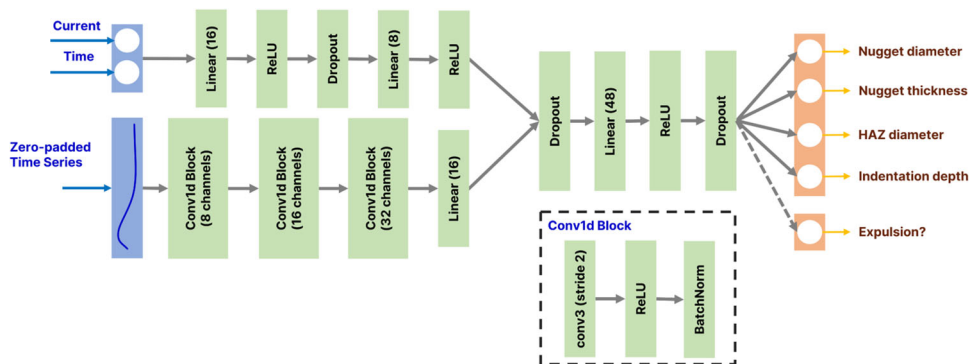
stride of 2 and kernel size of 3, ReLU activation, and batch normalization successively double the channels ($8 \diamond 16 \diamond 32$) while dividing the signal length by two. The time series is zero padded to ensure all inputs have the same length. A linear layer maps the outputs of the last CNN block into 16 features for downstream tasks. Current and time process parameters flow through the upper branch of the network to produce 8 features that are concatenated with the 16 convolutional features. A final series of fully connected layers maps these 24 features to the weld quality metrics or expulsion prediction.

Feature importance using Shapley values

While DL models achieve high accuracy and good performance on a wide array of tasks, the models themselves are black boxes that lack interpretability which would explain how the model arrived at a prediction. One helpful approach to improve interpretability is to determine how each input feature affects the model's output. For linear models that express the output as a weighted sum of input features, interpreting the features is straightforward since each feature's weight indicates its importance (Lundberg & Lee, 2017; Štrumbelj & Kononenko, 2014). In contrast, most highly-capable, nonlinear ML models have no straightforward interpretation. However, their interpretability can be improved by linearly approximating them for individual examples to produce local linear explanations (Štrumbelj & Kononenko, 2014). In other words, an interpretable approach could assign values to individual features that linearly sum to the model's output and represent how each feature contributed to the prediction. This can be achieved with Shapley values from cooperative game theory (Shapley, 1951).

A cooperative game consists of a set of players and a characteristic function that defines the value of any coalition (subset) of possible players. In ML the "players" are the features, and the characteristic function value is the model's prediction when trained on the features in the coalition. However, retraining the model on all coalitions of features is not feasible. Instead, the model's prediction with some features removed can be approximated by the expected value of the

Fig. 4 CNN-based architecture for predicting weld quality or detecting expulsions



model prediction across all values of the “removed” features in the training data (i.e., averaging out the removed features). Let x' be a binary vector of all ones with length $|x'|$ and $h_x(x')$ be a mapping such that $x = h_x(x')$. Let z' be a binary vector (with $|z'|$ nonzero elements) describing which features in x are “playing” in a coalition, and $f_x(z') = E[f(h_x(z'))|z'_S]$ be the expected value of the model across all remaining potential coalitions when the nonzero elements S of z' are fixed. Then, the SHAP definition of the Shapley value for model interpretation is

$$\phi_i(f, x) = \sum_{z' \subseteq x'} \frac{|z'|(|x'| - |z'| - 1)}{|x'|} [f_x(z') - f_x(z' \setminus i)] \tag{4}$$

where $z' \setminus i$ is the coalition with the i th feature removed (Lundberg & Lee, 2017). This value represents the mean change in model output when a feature is added to an existing coalition and the prediction is marginalized over the remaining excluded features. The value of an empty coalition is the mean model prediction across all training examples, and the local linear prediction becomes

$$f(x) = g(x') = \phi_0 + \sum_{i=1}^{|x'|} \phi_i x'_i \tag{5}$$

Unfortunately, computing the SHAP values in (5) from (4) remains computationally prohibitive for combinatoric reasons. Shapley sampling values, Kernel SHAP, and Deep LIFT offer three alternatives for efficiently estimating the Shapley values defined by SHAP (Štrumbelj & Kononenko, 2014; Lundberg & Lee, 2017). Kernel SHAP extends Local Interpretable Model-Agnostic Explanations (LIME) to produce proper Shapley values, and Deep LIFT estimates Shapley values via backpropagation through network layers (Ribeiro et al., 2016; Shrikumar et al., 2017). In this study, Shapley values associated with MLP feature-based modeling are estimated with Deep LIFT.

In Deep LIFT, a multiplier is the contribution of a difference Δx from a reference input x^0 to the difference in a specific target neuron’s output Δt from the reference output value, normalized by the input difference (i.e., a “finite gradient” of the target neuron with respect to the input):

$$m_{\Delta x \Delta t} = \frac{C_{\Delta x \Delta t}}{\Delta x} \tag{6}$$

If a single layer separates the input neuron and the target neuron, the multiplier can be written as

$$m_{\Delta x \Delta t} = \sum_i m_{\Delta x \Delta y_i} m_{\Delta y_i \Delta t} \tag{7}$$

which is the sum of multipliers across all possible paths from the input to the target neuron. This “chain rule for multipliers” enables input-to-output multipliers to be found through backpropagation (Shrikumar et al., 2017). To avoid asymmetries that might destroy information when multipliers are backpropagated, Deep LIFT proposes separate calculations of the positive and negative differences between the input and output. Consider a neural network operation $y = f(x)$ (e.g., linear layer, activation function, pooling, etc.) with an input difference $\Delta x = \Delta x^+ + \Delta x^-$ and output difference $\Delta y = \Delta y^+ + \Delta y^-$. The contributions of Δx^+ and Δx^- to Δy^+ and Δy^- can be written

$$\begin{aligned} \Delta y^+ &= \frac{1}{2} (f(x^0 + \Delta x^+) - f(x^0)) \\ &\quad + \frac{1}{2} (f(x^0 + \Delta x^- + \Delta x^+) - f(x^0 + \Delta x^-)) \\ \Delta y^- &= \frac{1}{2} (f(x^0 + \Delta x^-) - f(x^0)) \\ &\quad + \frac{1}{2} (f(x^0 + \Delta x^+ + \Delta x^-) - f(x^0 + \Delta x^+)) \end{aligned} \tag{8}$$

If Δx^+ and Δx^- are two “players” in a game with characteristic function $f(x)$, Δy^+ is computed as the Shapley value of Δx^+ since the change in model output is averaged across adding the Δx^+ to all possible coalitions (i.e., coalition #1 when Δx^- is not present and coalition #2 when Δx^-

is already present). Similarly, Δy^- functions as the Shapley value of Δx^- . The multipliers for the function $f(x)$ can be obtained via

$$m_{\Delta x^+ \Delta y^+} = \frac{\Delta y^+}{\Delta x^+}; m_{\Delta x^- \Delta y^-} = \frac{\Delta y^-}{\Delta x^-} \quad (9)$$

These can be backpropagated from the network outputs to inputs to compute the multipliers of the input features. The Shapley values for input feature x_i with respect to network output y can then be approximated as:

$$\phi_{x_i} = \phi_{x_i^+} + \phi_{x_i^-} = m_{\Delta x_i^+ \Delta y^+} \Delta x_i^+ + m_{\Delta x_i^- \Delta y^-} \Delta x_i^- \quad (10)$$

If the features can be approximated as independent and the model is approximated as linear, the reference input x^0 should be $E[x]$ to estimate SHAP values consistent with (7) (Lundberg & Lee, 2017). Thus, Deep LIFT can estimate the SHAP definition of Shapley values efficiently, providing explanations for model predictions. Importantly, the values depend on both the model and the data, so care should be taken when drawing broad conclusions—they explain the contributions of features present in the data to the model's predictions but are not necessarily reflective of any physical causation underlying the studied process. While they are primarily a local linear approximation technique, averaging SHAP value magnitudes across the data set can reveal the mean contributions of features for the problem at hand and offer a more interpretable assessment of complex NN models.

Experiments

To evaluate and compare MLP and CNN for RSW weld quality prediction and expulsion detection, weld experiments were conducted under different sheet fit-up conditions and process parameters to generate both sensing data and weld quality labels. In the experimental setup, a FANUC R-2000iC robotic arm held a CENTERLINE RSW gun driven by a WT6000s Medium Frequency DC weld controller with a 1 kHz inverter frequency. The gun contained two copper zirconium (CuZr) C15000 electrodes with 6-mm tips that were redressed every 15 welds to maintain their geometry. Experiments consisted of 113 trials: 42 trials used a constant welding time of 150 ms for welding currents ranging from 4 to 9 kA, and the remaining 71 trials tested welding times from 10 to 150 ms with currents of 5 kA and 8 kA, as shown in Table 2. All trials were performed on two 150 mm by 38 mm by 0.8 mm sheets of bare DP590 steel. Multiple sensors collected current, secondary voltage, electrode force, and electrode displacement data through a ZHOUFAN weld monitoring system that produced 1-ms averages of 500-kSPS

sensing data. Attached to the electrode shank, a MEATROL Rogowski coil rated at 0.5% accuracy measured the welding current, and voltage was measured with probes across the upper and lower electrode holders. A HEIDENHAIN linear encoder with 0.5- μ m accuracy collected displacement data, and a Kistler surface strain sensor (2% accuracy) monitored the welding force. Data was collected during both the active welding time and the subsequent holding period, and nugget thickness, nugget diameter, HAZ diameter, and indentation depth were labelled after the weld was completed (Table 2).

The MLP and CNN from Figs. 3 and 4, respectively, are trained on the experimental data to detect expulsions and predict the weld quality metrics (i.e., nugget diameter, nugget thickness, HAZ diameter, and indentation depth). Four training data variations evaluate how different data perspectives impact the performance of both models: training with and without force and displacement information provides insight into the model's capability with different sensor inputs, and both scenarios are run with and without trials with less than or equal to 60 ms of time series data to assess the impact of including partially formed nuggets. The choice of 60 ms was due to it being the approximate time when measured force and displacement stabilize, as discussed on Fig. 1. Thus, the four data scenarios are All Sensing Data (ALL), All Sensing Data from Trials > 60 ms (ALL-60), Current and Resistance Data (CR), and Current and Resistance Data from Trials > 60 ms (CR-60). Since very few expulsions occur in the data set, expulsion detection training applies 10x oversampling to the minority (expulsion) class to mitigate the negative impact of imbalance between expulsion and non-expulsion samples.

Data preprocessing normalizes the process parameters, features, and time series data to the range [0, 1]. When training the CNN, all signals are zero padded to the longest length in the data set since the length of the raw time series depends on the welding time. Each sensor forms one channel of the CNN input. Both the MLP and CNN use the MSE loss function for predicting weld quality values and BCE when detecting expulsions. Five-fold cross validation evaluates the model performance considering the limited quantity of data, and five random seeds capture variability in model initialization. All experiments are optimized with Adam (learning rate of 0.001 for MLP and 0.0001 for CNN) on a workstation with an NVIDIA GPU. For weld quality prediction, the CNN trainer uses early stopping on the validation loss to prevent overfitting with a maximum of 1500 passes through the data, while the MLP training runs for a constant 3000 epochs to avoid local minimums.

Results and discussion

Repeating the four data combinations and five folds across five random seeds generates 100 trained models for weld

Table 2 Process parameters and trials for multi-condition welding experiments

Condition	Force (kN)	Current (kA)	Time (ms)	Count
Standard	2.6	4–9	10–150	25
Edge proximity 3 mm	2.6	4–9	10–150	26
Edge proximity 6 mm	2.6	4–9	150	6
Initial gap 1 mm	2.6	4–9	150	6
Initial gap 2 mm	2.6	4–9	10–150	24
Off normal 3°	2.6	4–9	150	6
Off normal 6°	2.6	4–9	20–150	20

quality prediction and 100 trained models for expulsion detection for both MLP and CNN architectures.

MLP feature-based prediction versus CNN raw sensing-based prediction

The ALL and ALL-60 columns of Tables 3 and 4 show the MLP and CNN performance with the full suite of available sensing data. Twenty-five values—five folds repeated for each of five seeds—provide mean and standard deviation information for each entry. Performance metrics include the coefficient of determination R^2 , Root Mean Squared Error (RMSE), and relative error (RMSE scaled by the mean target value).

Both the MLP and CNN achieved a best relative error below 16% when using all the sensing data (ALL or ALL-60). Figure 5a and b present the ALL and ALL-60 relative error as the first two bars in each variable group for MLP and CNN, respectively. When using all the welding experiments and sensing data (ALL), the MLP model achieves slightly lower relative error when compared to excluding the trials shorter than 60 ms (ALL-60). This may indicate that the selected features remain robust and informative even for partially incomplete welds. Furthermore, the MLP offers approximately a 5-point advantage in relative error over the CNN for nugget geometry when using ALL. This also supports the idea that the MLP features may be more robust due to their purposeful design from a physical understanding of the RSW sensing curves shown in Fig. 1.

Interestingly, while recurrent neural networks like Long Short-Term Memory (LSTM) are often suitable for time series sensing data, exploratory experiments with a bidirectional, three-unit LSTM demonstrated significantly worse performance than both the MLP and CNN. Despite including memory gates, the LSTM still may be struggling to remember information from throughout the sequence since key points could occur at different times in different experiments. Future work could incorporate an attention mechanism to better highlight the sequence elements most relevant to the quality predictions.

In contrast, the CNN appears to struggle to learn good features when the short trials are present (ALL), resulting in worse performance than the MLP. The short trials increase the variation in the time series training data and could make it harder for the CNN to identify consistent trends in the data and learn features automatically. Additionally, since the time series are zero-added, the short-duration input signals contain mostly zeros. Thus, the valid information from the start of the signal could be quickly attenuated through the successive CNN layers and surrounded by “features” extracted from the zero padding. However, when the problematic short trials are removed, the CNN model noticeably improves and reduces the relative error on nugget geometry by more than 5 points and HAZ relative error by 1.1 points. This supports the conclusion that CNN can perform well when the quality and variation of the input data are consistent enough for automatic feature learning.

Figure 6 shows the expulsion detection validation accuracies for the MLP and CNN models for ALL and ALL-60 as the first two bar groups. With all the welding experiments and sensing data (ALL), both the MLP model and CNN reach higher than 95% accuracy, and the CNN achieves 100% detection. Removing the 60 ms and shorter trials diminish the accuracy. With those trials excluded, the relative rate of expulsions in the data set increases since no expulsions occurred in trials less than 100 ms in length. Therefore, this increase in relative expulsion frequency could accentuate the misclassification error for the MLP model that was seen for ALL, reducing its accuracy on ALL-60. The CNN also experiences reduced accuracy for ALL-60, but its 100% accuracy for ALL indicates that this may not be related to the relative frequency of expulsions in the training data. Instead, it could be caused by the variations in the splitting of folds for training and validation. The data set is already limited, so removing the short trials could produce new training folds that contain fewer expulsion examples to learn from, hurting the overall performance. Future work should explore additional methods for mitigating the impact of limited expulsion examples and improving detection accuracy.

Table 3 Weld quality prediction and expulsion detection results using MLP for different combinations of data

Quantity	Validation metric	MLP			
		ALL	ALL-60	CR	CR-60
Nugget diameter	R ²	0.924 ± 0.042	0.862 ± 0.086	0.906 ± 0.069	0.876 ± 0.082
	RMSE (mm)	0.451 ± 0.128	0.530 ± 0.203	0.503 ± 0.158	0.488 ± 0.161
	RMSE / Mean (%)	15.0 ± 4.6	15.6 ± 7.9	16.8 ± 5.8	14.1 ± 5.9
Nugget thickness	R ²	0.912 ± 0.053	0.818 ± 0.114	0.840 ± 0.085	0.745 ± 0.160
	RMSE (mm)	0.130 ± 0.034	0.160 ± 0.056	0.181 ± 0.045	0.188 ± 0.060
	RMSE / Mean (%)	15.9 ± 4.0	18.3 ± 9.2	22.4 ± 6.2	20.9 ± 8.4
HAZ diameter	R ²	0.939 ± 0.023	0.894 ± 0.085	0.952 ± 0.026	0.914 ± 0.047
	RMSE (mm)	0.276 ± 0.058	0.292 ± 0.115	0.246 ± 0.081	0.276 ± 0.086
	RMSE / Mean (%)	6.1 ± 1.3	5.9 ± 2.4	5.4 ± 1.9	5.6 ± 2.0
Indentation depth	R ²	0.802 ± 0.270	0.848 ± 0.193	− 0.488 ± 2.622	− 0.367 ± 1.489
	RMSE (mm)	0.026 ± 0.006	0.029 ± 0.014	0.071 ± 0.028	0.085 ± 0.025
	RMSE / Mean (%)	1.6 ± 0.4	1.8 ± 0.9	4.5 ± 1.8	5.4 ± 1.7
Expulsion	Accuracy (%)	96.0 ± 8.2	93.9 ± 12.4	75.3 ± 11.4	64.2 ± 20.7

Bolded values are the best performance for a given quantity

All sensing data versus current and resistance data

In production plants, collecting force and displacement data is restricted by cost and the technical difficulty of widespread installation and consistent calibration in the factory environment. Thus, current and DR may be the only information available for predictive models. The final two columns of Tables 3 and 4 record the MLP and CNN performance metrics when excluding force and displacement sensing data from training (CR and CR-60). While the relative error increases, both the MLP and CNN can still reach relative errors below 21%. Figure 5 also includes the CR and CR-60 results for comparison with the ALL and ALL-60 results. Nugget geometry predictions suffer by about 2 to 6 points when the MLP is restricted to CR data, although excluding the shorter welds (CR-60) improves relative error for both nugget diameter and nugget thickness. Indentation depth suffers a clear increase in relative error for both CR and CR-60, likely attributable to the highly informative nature of the excluded force and displacement data for this quantity. With respect to HAZ diameter prediction, the MLP model trained on CR nearly

matches the performance for ALL and ALL-60, and this may be because HAZ diameter predictions are less impacted by excluding the axial perspective offered by force and displacement.

The CNN experiences similar outcomes, with nugget geometry predictions displaying increased relative error while HAZ diameter predictions show less variation. As with the MLP, relative error for indentation depth increases since the salient information from force and displacement signals is no longer present. However, while both the MLP and CNN experience reduced performance without force and displacement data, the differences are often slight, especially for the CNN. This indicates that current and resistance alone may provide enough information for adequately low relative error when predicting weld quality metrics.

SHAP value analysis of feature contribution

Analyzing the MLP feature contributions with Shapley values provides additional insight into the correlations learned by the MLP. Figure 7 plots the top three Shapley values for the

Table 4 Weld quality prediction and expulsion detection results using CNN for different combinations of data

Quantity	Validation metric	CNN			
		ALL	ALL-60	CR	CR-60
Nugget diameter	R ²	0.858 ± 0.120	0.887 ± 0.055	0.859 ± 0.143	0.871 ± 0.089
	RMSE (mm)	0.588 ± 0.173	0.475 ± 0.106	0.582 ± 0.167	0.493 ± 0.149
	RMSE / Mean (%)	19.7 ± 6.2	13.6 ± 4.0	19.4 ± 5.0	14.1 ± 5.5
Nugget thickness	R ²	0.849 ± 0.106	0.860 ± 0.089	0.842 ± 0.091	0.821 ± 0.123
	RMSE (mm)	0.166 ± 0.053	0.134 ± 0.038	0.173 ± 0.042	0.151 ± 0.037
	RMSE / Mean (%)	20.8 ± 7.6	14.9 ± 5.6	21.6 ± 6.5	16.7 ± 5.2
HAZ diameter	R ²	0.854 ± 0.077	0.818 ± 0.107	0.839 ± 0.082	0.826 ± 0.077
	RMSE (mm)	0.414 ± 0.104	0.398 ± 0.117	0.438 ± 0.112	0.394 ± 0.103
	RMSE / Mean (%)	9.1 ± 2.4	8.0 ± 2.6	9.6 ± 2.6	7.9 ± 2.3
Indentation depth	R ²	0.792 ± 0.319	0.813 ± 0.238	0.127 ± 0.716	-0.145 ± 1.188
	RMSE (mm)	0.028 ± 0.011	0.032 ± 0.016	0.069 ± 0.030	0.084 ± 0.036
	RMSE / Mean (%)	1.8 ± 0.7	2.0 ± 1.0	4.3 ± 1.9	5.4 ± 2.3
Expulsion	Accuracy (%)	100.0 ± 0.0	97.8 ± 7.5	82.9 ± 17.4	78.6 ± 22.5

Bolded values are the best performance for a given quantity

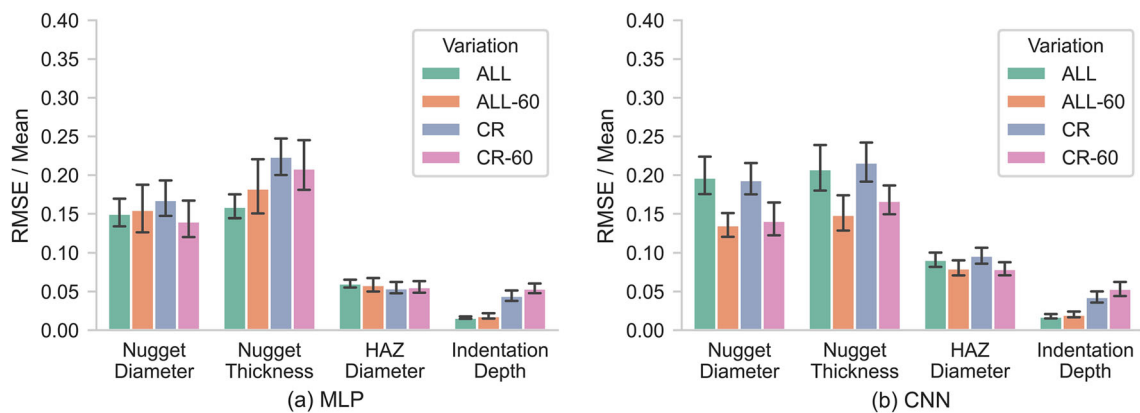


Fig. 5 Weld quality prediction relative error (RMSE scaled by the mean target value) for **a** MLP and **b** CNN

MLP when using all features to predict weld quality metrics (ALL). The SHAP method attributes high average contributions to Displacement ΔCD (i.e., across the holding period) for nugget geometry metrics. During this period, the displacement sensor detects how the metal contracts during cooling, providing information about the final size of the nugget. The Shapley values also show that the average measured current

has a higher contribution to nugget diameter than to nugget thickness. Current would be more informative for predicting the diameter versus thickness since it provides insight into the heat flux entering the weld that affects the radial growth of the nugget; axial growth can be better predicted from data that more directly relates to nugget thickness such

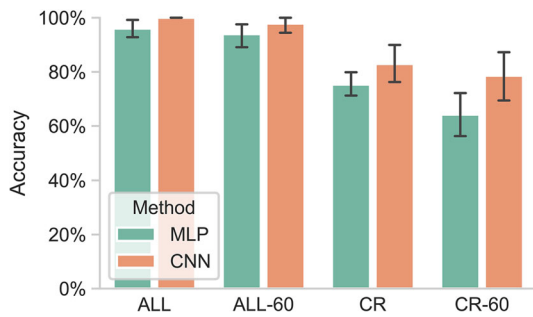


Fig. 6 Expulsion prediction accuracies for MLP and CNN models across the data scenarios

as force (e.g., Force ΔAB) and displacement (e.g., Displacement ΔCD). In addition, Fig. 7 indicates that HAZ diameter predictions are most affected by input current, average of measured current, and time. This can be understood since current influences the heat flux, and time controls how long that heat is applied. Therefore, these variables together correlate to the total energy injected into the weld to create the HAZ.

When short trials are removed, SHAP analysis attributes increased impact to the current features (see Fig. 8). Input current and average of measured current remain roughly constant since current is an RSW process parameter. Therefore, when all trials are included (Fig. 7), these constant values are paired with many different nugget geometries as the growth is halted at various points. The lack of one-to-one mapping means that current cannot resolve the geometry of a partially formed nugget. However, once the nugget has enough time to approach its terminal geometry, the RSW current parameters correlate well with this final size, allowing them to contribute

more to the prediction in the ALL-60 case. Displacement ΔCD retains a slightly above average contribution to the nugget thickness but is not considered as influential given the more well-defined relationship between nugget geometry and average of measured current when short trials are removed. As for ALL, both time and current play an important role in predicting the HAZ diameter and nugget diameter, as the radial diameter growth correlates better with the total input energy determined by time and current than the axially focused force and displacement data.

Figure 9 plots the top three mean Shapley values for MLP features when trained on only current and resistance sensing data and process parameters (i.e., CR). Resistance ΔAB significantly contributes to predicting nugget thickness, which can be understood since resistance during this melting and heating phase will be dependent on the axial thickness of the nugget. The Shapley value for Resistance ΔBC indicates that it contributes most to predicting the indentation depth. This follows an understanding that decreasing resistance at the end of the molten period will be correlated with the total contact area between the electrode and sheet metal, which increases as indentation depth increases. As with ALL-60, time and current features combine for the largest contribution to nugget diameter, and the total energy information carried by these parameters appears to contribute most to both radial nugget growth and HAZ diameter.

Removing shorter trials from the MLP training produces the Shapley values shown in Fig. 10 for the CR-60 case. The contribution of time is diminished since the durations are more uniform, and thus time no longer plays as crucial a role in predicting the nugget size. However, current remains a significant contributor to HAZ diameter in keeping with

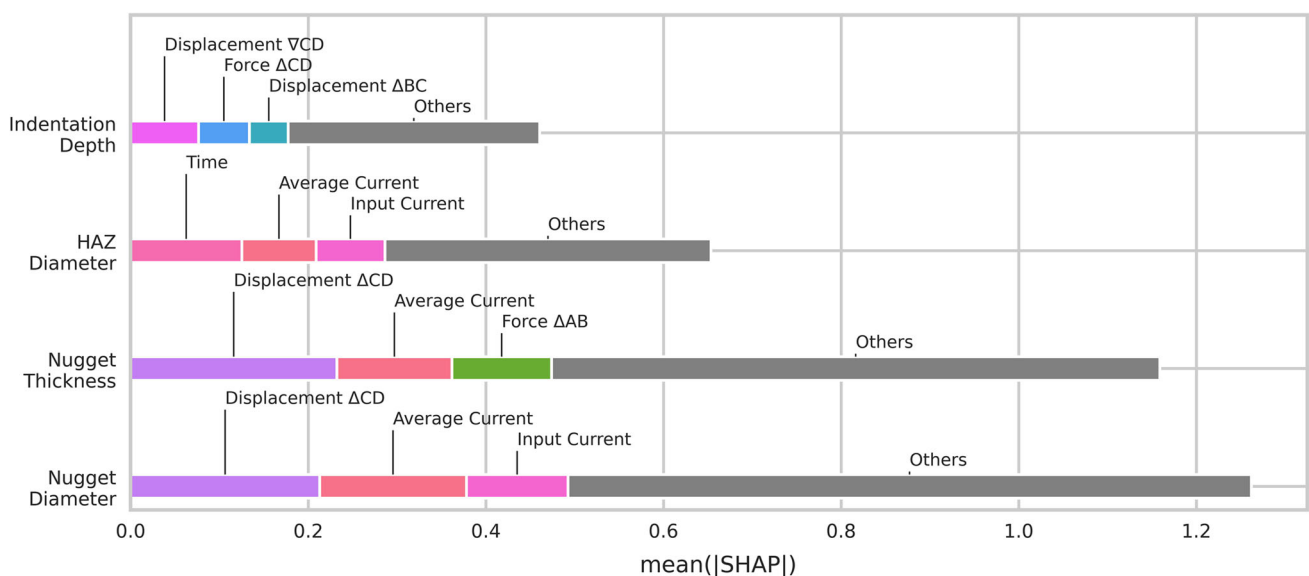


Fig. 7 Top three mean SHAP value magnitudes for MLP features when using ALL data

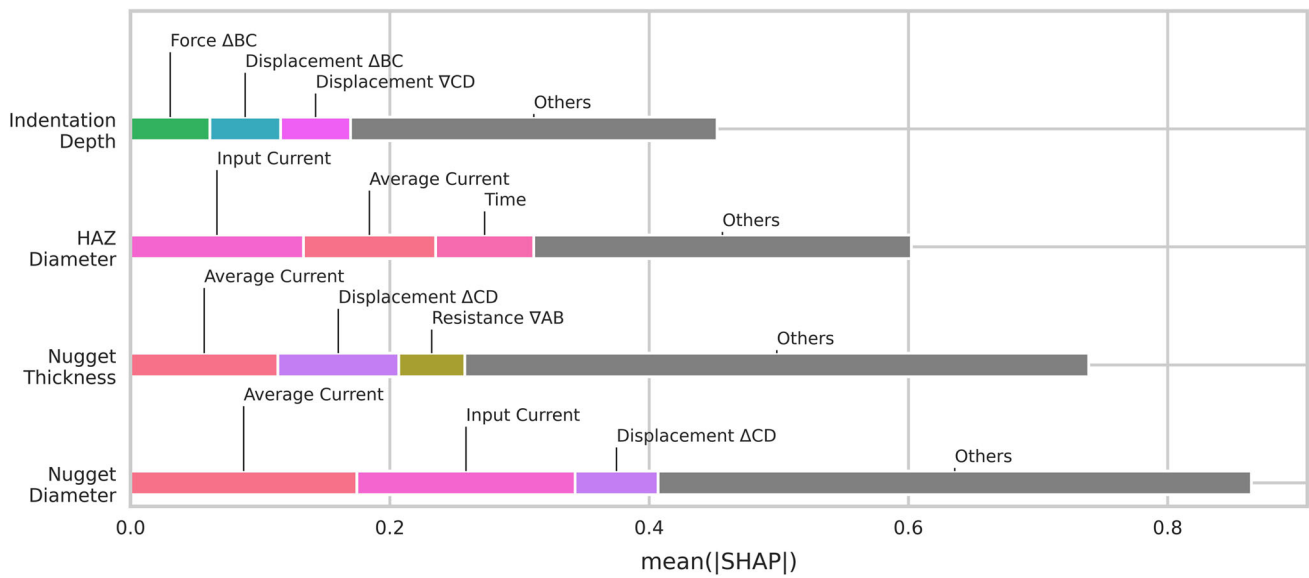


Fig. 8 Top three mean SHAP value magnitudes for MLP features when using ALL-60 data

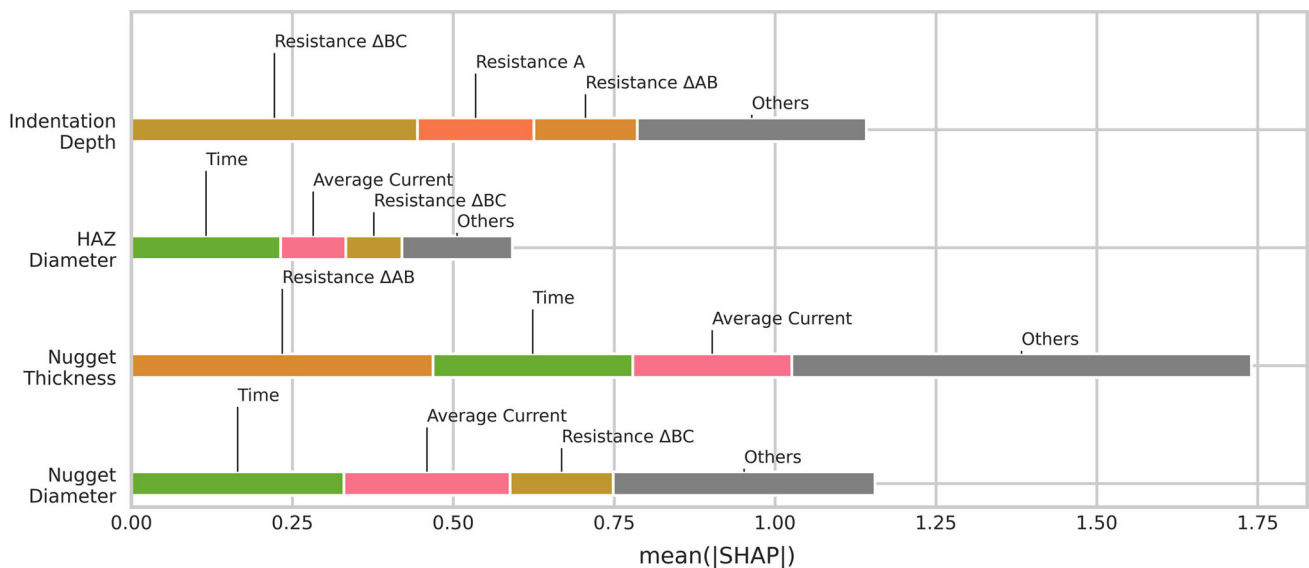


Fig. 9 Top three mean SHAP value magnitudes for MLP features when using CR data

the ALL, ALL-60, and CR cases, and factors heavily in predicting the non-axial metric of nugget diameter. In contrast, earlier parts of the resistance curve (e.g., Resistance A and ΔAB) show larger Shapley values for the related metrics of nugget thickness and indentation depth than in previous scenarios. Longer duration trials will have more fully formed nuggets, and therefore early parts of the sensing curves can be useful without the possibility of the weld being terminated prematurely.

This SHAP value analysis provides insight into how the trained MLP model perceives the contribution of input features to the output prediction. While the results can broadly inform feature selection in future models, the primary value

of SHAP analysis is verifying that the features deemed important by the model mirror the empirical physical understanding of RSW process fundamentals. Dependent on both the model architecture and training data, the SHAP values provide an interpretable explanation to ensure that the tested MLP model learned physically meaningful relationships and provide broad insights into feature impact that can influence future studies on RSW process optimization. Further analysis with additional welding experiments and model architectures should be performed to solidify any generalizations broadly connecting specific features to weld quality metrics.

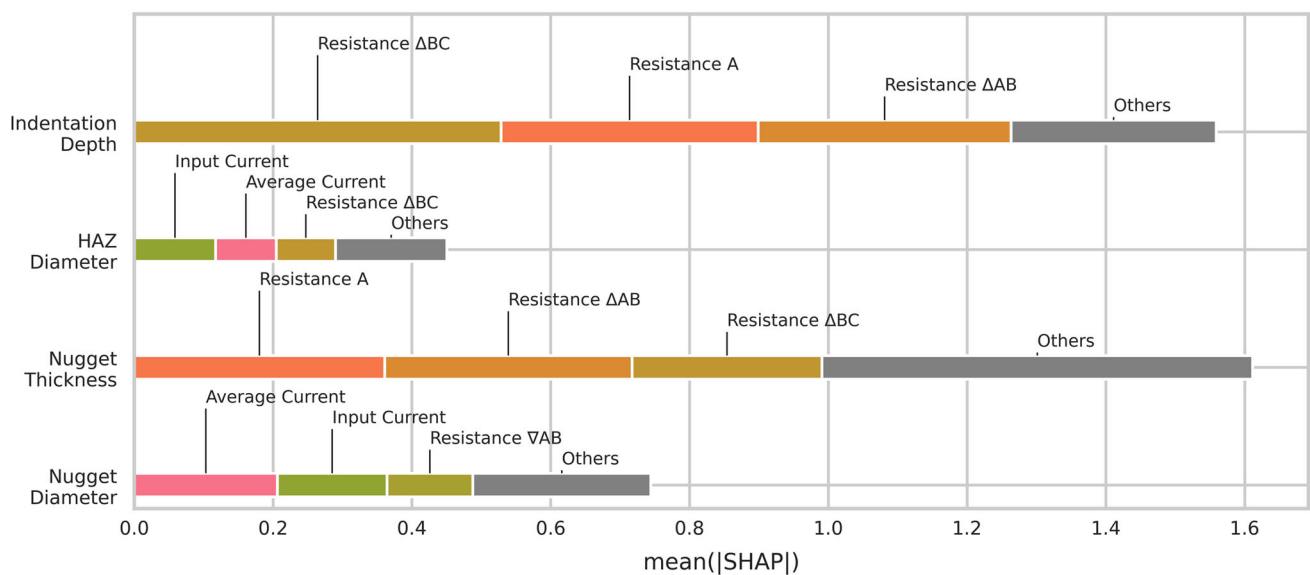


Fig. 10 Top three mean SHAP value magnitudes for MLP features when using CR-60 data

Conclusion

This study evaluates and compares the performance of a feature-based MLP and raw sensing-based CNN when predicting RSW weld quality metrics and detecting expulsions. The models are trained with two types of inputs—all sensing (i.e., using all current, DR, force, and displacement signals) and with only current and resistance signals—to determine the impact of different sensing combinations on prediction performance. Both the MLP and CNN can perform at a similar level when predicting nugget metrics and expulsions. The MLP achieves 14.1% relative error on nugget diameter prediction, 15.9% on nugget thickness, 5.4% on HAZ diameter, and 1.6% on indentation depth; the CNN reaches 13.6% relative error on nugget diameter prediction, 14.9% on nugget thickness, 7.9% on HAZ diameter, and 1.8% on indentation depth. Excluding force and displacement signals slightly reduces the accuracies but still generates comparable prediction performance. The CNN model could achieve the lowest relative error when shorter trials with incomplete nugget formation were removed, indicating that data quality and consistency are critical for automatic feature learning methods. Analyzing Shapley values for the MLP model reveals the contributions from each input feature to the model's output and can be validated against the physical understanding of RSW process dynamics. For example, displacement from the holding period contributes significantly to nugget geometry prediction since it reflects how the metal in the weld region contracts to its final size. Future work should explore additional architectural tuning that could improve model performance, explore ways to translate lab-developed models effectively to production environments, and leverage these

models to optimize RSW processes and improve quality consistency.

References

- Ao, S., Li, C., Huang, Y., & Luo, Z. (2020). Determination of residual stress in resistance spot-welded joint by a novel X-ray diffraction. *Measurement*, *161*, 107892. <https://doi.org/10.1016/j.measurement.2020.107892>
- Batista, M., Furlanetto, V., & Duarte Brandi, S. (2020). Analysis of the behavior of dynamic resistance, electrical energy and force between the electrodes in resistance spot welding using additive manufacturing. *Metals*, *10*(5), 690. <https://doi.org/10.3390/met10050690>.
- Chen, J., Feng, Z., Wang, H. P., Carlson, B. E., Brown, T., & Sigler, D. (2018). Multi-scale mechanical modeling of al-steel resistance spot welds. *Materials Science and Engineering: A*, *735*, 145–153. <https://doi.org/10.1016/j.msea.2018.08.039>.
- El-Sari, B., Biegler, M., & Rethmeier, M. (2021). Investigation of the extrapolation capability of an artificial neural network algorithm in combination with process signals in resistance spot welding of advanced high-strength steels. *Metals*, *11*(11), 1874. <https://doi.org/10.3390/met11111874>.
- Hwang, I., Yun, H., Yoon, J., Kang, M., Kim, D., & Kim, Y. M. (2018). Prediction of resistance spot weld quality of 780 MPA grade steel using adaptive resonance theory artificial neural networks. *Metals*, *8*(6), 453. <https://doi.org/10.3390/met8060453>.
- Ji, C. T., & Zhou, Y. (2004). Dynamic Electrode Force and displacement in resistance spot welding of aluminum. *Journal of Manufacturing Science and Engineering*, *126*(3), 605–610. <https://doi.org/10.1115/1.1765140>
- Lee, J., Noh, I., Jeong, S. I., Lee, Y., & Lee, S. W. (2020). Development of real-time diagnosis framework for angular misalignment of robot spot-welding system based on machine learning. *Procedia Manufacturing*, *48*, 1009–1019. <https://doi.org/10.1016/j.promfg.2020.05.140>.
- Lundberg, S., & Lee, S-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. Von Luxburg, S. Bengio,

- H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems*, 30. <https://doi.org/10.48550/arXiv.1705.07874>
- Ma, Y., Wu, P., Xuan, C., Zhang, Y., & Su, H. (2013). Review on techniques for on-line monitoring of resistance spot welding process. *Advances in Materials Science and Engineering*,. <https://doi.org/10.1155/2013/630984>
- Ribeiro, M. T., Sing, S., & Guestrin, C. (2016). “Why should I trust you?” Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135–1144). <https://doi.org/10.48550/arXiv.1602.04938>
- Shapley, L. S. (1951, August 21). Notes on the n-person game—II: The value of an n-person game [Memorandum]. RAND Corporation. Retrieved from https://www.rand.org/content/dam/rand/pubs/research_memoranda/2008/RM670.pdf
- Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning important features through propagating activation differences. In *Proceedings of the 34th international conference on machine learning* (Vol. 70, pp. 3145–3153). <https://doi.org/10.48550/arXiv.1704.02685>
- Štrumbelj, E., & Kononenko, I. (2014). Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems*, 41, 647–665. <https://doi.org/10.1007/s10115-013-0679-x>.
- Vignesh, K., Perumal, A. E., & Velmurugan, P. (2019). Resistance spot welding of AISI-316L SS and 2205 DSS for predicting parametric influences on weld strength—Experimental and FEM approach. *Archives of Civil and Mechanical Engineering*, 19(4), 1029–1042. <https://doi.org/10.1016/j.acme.2019.05.002>.
- Wang, S. C., & Wei, P. S. (2000). Modeling dynamic electrical resistance during resistance spot welding. *Journal of Heat Transfer*, 123(3), 576–585. <https://doi.org/10.1115/1.1370502>.
- Wang, X. J., Zhou, J. H., Yan, H. C., & Pang, C. K. (2017). Quality monitoring of spot welding with advanced signal processing and data-driven techniques. *Transactions of the Institute of Measurement and Control*, 40(7), 2291–2302. <https://doi.org/10.1177/0142331217700703>
- Wang, Y., Rao, Z., & Wang, F. (2020). Heat evolution and nugget formation of resistance spot welding under multi-pulsed current waveforms. *The International Journal of Advanced Manufacturing Technology*, 111(11–12), 3583–3595. <https://doi.org/10.1007/s00170-020-06337-z>.
- Wu, N., Chen, S., & Xiao, J. (2018). Wavelet analysis-based expulsion identification in electrode force sensing of resistance spot welding. *Welding in the World*, 62(4), 729–736. <https://doi.org/10.1007/s40194-018-0594-6>.
- Xia, Y. J., Zhou, L., Shen, Y., Wegner, D. M., Haselhuhn, A. S., Li, Y. B., & Carlson, B. E. (2021). Online measurement of Weld Penetration in robotic resistance spot welding using electrode displacement signals. *Measurement*, 168, 108397. <https://doi.org/10.1016/j.measurement.2020.108397>.
- Xia, Y. J., Shen, Y., Zhou, L., & Li, Y. B. (2020). Expulsion intensity monitoring and modeling in resistance spot welding based on electrode displacement signals. *Journal of Manufacturing Science and Engineering*, 143(3). <https://doi.org/10.1115/1.4048441>.
- Xing, B., Xiao, Y., & Qin, Q. H. (2018). Characteristics of shunting effect in resistance spot welding in mild steel based on electrode displacement. *Measurement*, 115, 233–242. <https://doi.org/10.1016/j.measurement.2017.10.049>.
- Xing, B., Xiao, Y., Qin, Q. H., & Cui, H. (2017). Quality Assessment of resistance spot welding process based on dynamic resistance signal and random forest based. *The International Journal of Advanced Manufacturing Technology*, 94(1–4), 327–339. <https://doi.org/10.1007/s00170-017-0889-6>.
- Zaharuddin, M. F., Kim, D., & Rhee, S. (2017). An ANFIS based approach for predicting the weld strength of resistance spot welding in Artificial Intelligence Development. *Journal of Mechanical Science and Technology*, 31(11), 5467–5476. <https://doi.org/10.1007/s12206-017-1041-0>.
- Zhao, D., Bezgans, Y., Wang, Y., Du, W., & Lodkov, D. (2020a). Performances of dimension reduction techniques for welding quality prediction based on the dynamic resistance signal. *Journal of Manufacturing Processes*, 58, 335–343. <https://doi.org/10.1016/j.jmapro.2020.08.037>.
- Zhao, D., Bezgans, Y., Wang, Y., Du, W., & Vdonin, N. (2021). Research on the correlation between dynamic resistance and quality estimation of resistance spot welding. *Measurement*, 168, 108299. <https://doi.org/10.1016/j.measurement.2020.108299>.
- Zhao, D., Ivanov, M., Wang, Y., & Du, W. (2020b). Welding quality evaluation of resistance spot welding based on a hybrid approach. *Journal of Intelligent Manufacturing*, 32(7), 1819–1832. <https://doi.org/10.1007/s10845-020-01627-5>.
- Zhou, L., Xia, Y. J., Shen, Y., Haselhuhn, A. S., Wegner, D. M., Li, Y. B., & Carlson, B. E. (2021). Comparative study on resistance and displacement based adaptive output tracking control strategies for resistance spot welding. *Journal of Manufacturing Processes*, 63, 98–108. <https://doi.org/10.1016/j.jmapro.2020.03.061>.
- Zhou, L., Zheng, W., Li, T., Zhang, T., Zhang, Z., Zhang, Y., Wu, Z., Lei, Z., Wu, L., & Zhu, S. (2020). A material stack-up combination identification method for resistance spot welding based on dynamic resistance. *Journal of Manufacturing Processes*, 56, 796–805. <https://doi.org/10.1016/j.jmapro.2020.04.051>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.