

Efficient and Generalizable Machine Learning for Inline Defect Detection in Battery Laser Welding

Xijia Zhao¹, Joseph Kershaw², Masoud Pour³, Junjie Ma³, Hassan Ghassemi-Armaki³, Blair Calson³, and Peng Wang^{1,2,*}

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

*Corresponding Author: Edward.Wang@uky.edu

Abstract:

Laser welding, characterized by its high energy density, precision, and automation compatibility, has been extensively applied in electrical vehicle battery manufacturing. However, inline process monitoring and quality assurance in laser welding remains a challenge, due to the brief process yet intense energy output. Upon inline Laser Welding Monitoring (LWM) sensing that encompasses plasma, thermal, and back reflection measurement, this study presents an applicable and generalizable machine learning framework for efficient and effective sensing data analytics and extraction of physically meaningful features, to detect system configuration and welding quality defects. The framework first establishes a comprehensive feature pool from time and frequency domains and employs Shapley Additive Explanations (SHAP) and Principal Component Analysis (PCA) to identify a minimal but effective feature subset for defect detection. Subsequently, upon the identified feature set, separate Multi-Layer Perceptron (MLP) models are developed to distinguish system configuration and quality defects. Experimental results demonstrate a 99% accuracy of this framework in defect detection. Attributed to its high computational efficiency, this approach is suitable for real-time decision-making in practical deployment, alongside its physical explainability and generalizability.

Keywords : Measurement; Laser Welding; Machine Learning

1. Introduction

The growing market of Electric Vehicles (EVs), attributed to increased environmental awareness and supportive governmental policies, drives the development and breakthroughs of EV manufacturing technologies. As a critical component in EVs, battery cell and module production and reliability are pivotal in determining the vehicle's range, efficiency, and overall performance. Battery production involves a series of intricate manufacturing processes, among which the laser welding process that joins battery cells and modules gains the most attention, considering its high energy density, precision, and compatibility with automation (Feng et al., 2023).

Laser welding offers several advantages over other welding techniques, including reduced thermal stress which is crucial for maintaining the integrity and functionality of the battery cells. However, inline process monitoring and quality assurance in laser welding remains a challenge, due to the brief process yet intense energy output. Laser Welding Monitoring (LWM) has recently emerged as a promising sensing technique that can keep pace with the rapid dynamics and provide immediate feedback (Sadeghian et al., 2022). LWMs primarily measure the attributes that can reflect the intricate

process dynamics, e.g., light emission from plasma, signal of vapor plume, and information of molten pool (Cai et al., 2020). Correspondingly, photodiode sensors are leveraged to independently capture light emission of plasma, thermal, and laser reflection (Eriksson et al., 2009). LWM data analysis contributes to various welding quality prediction, defect detection, and process control tasks, such as optimizing welding parameters (Angeloni et al., 2024), predicting weld penetration depth (Li et al., 2022), and detecting weld defects (Zhang et al., 2019).

However, because of the lack of deep understanding of laser welding process dynamics and its correlations with LWM signal characteristics, most prior studies adopt a data-driven Machine Learning (ML)-based approach to analyze LWM data. She et al. (2024) established a neural network model to correlate features from plasma spectrum and molten pool images to prediction of penetration depth in titanium alloy welding. Brežan et al. (2023) built a random forest model to classify LWM signals against good or bad welds. Kim et al. (2022) employed a deep learning model to predict laser-beam absorptance inside the keyhole. These data-driven approaches bypass signal characterization and feature extraction by directly feeding raw

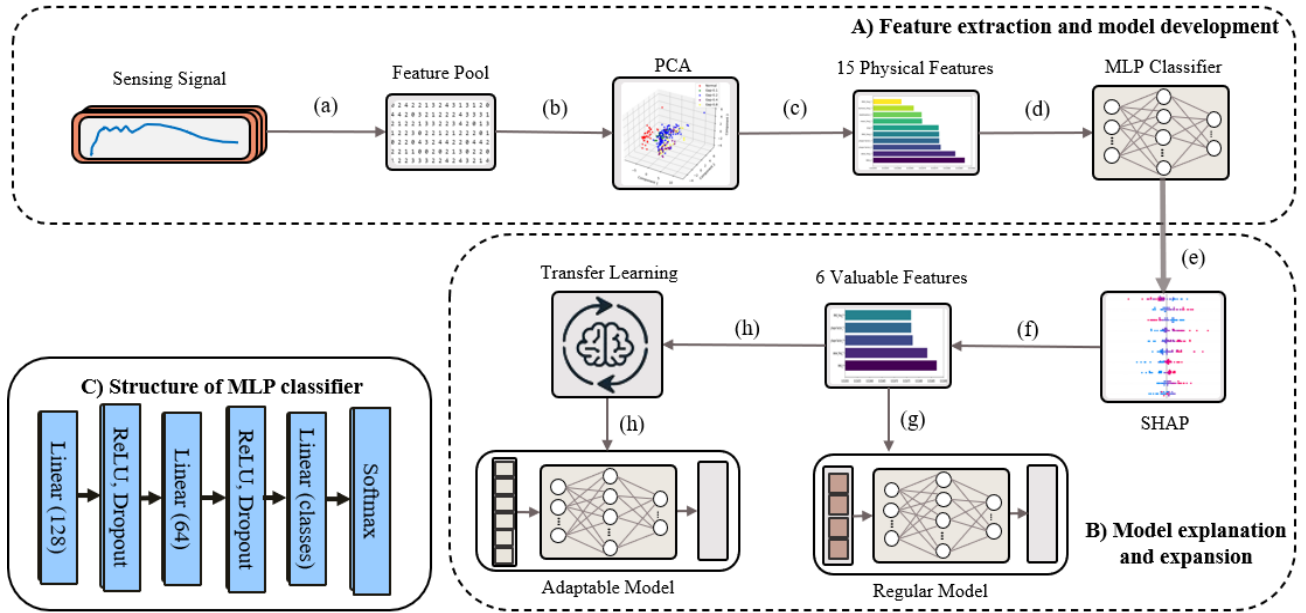


Fig.1 Overview of the developed explainable and generalizable ML framework for welding system defect detection in laser welding: **(A)** feature extraction and classifier development; **(B)** model explanation and generalization. In subtask **(A)**, step **(a)** represents the extraction of physically meaningful features from LWM sensing signal, step **(b)** refers to correlation and PCA analysis for reducing feature dimensionality, leading to 15 features selected in step **(c)**, followed by step **(d)** of building a MLP classifier upon the selected features. In subtask **(B)**, steps **(e)** and **(f)** represent SHAP analysis for further refinement of the feature set based on features' contributions to model prediction results, and step **(g)** represents refining the MLP model based on finally selected 6 features. Step **(h)** indicates the transfer learning-based generalization of the MLP model with the 6 features when defect categories increase.

data into ML models, while demonstrating good performance, struggling with decision-making explainability and model generalizability.

Prior studies also tried to extract physically meaningful statistical features from sensing signals and develop feature-based ML models for quality prediction and defect detection. Will et al. (2022) investigated a feature set efficient for spatter detection. Stavropoulos et al. (2022) developed two feature-based ML models to identify bad welds. Chianese et al. (2022) extracted a set of time-frequency domain features as inputs to ML models to detect variations in weld penetration depth and part-to-part gap.

However, the methods of feature extraction in previous studies which partly include both time-domain features (e.g., average, standard deviation, maximum) and time-frequency domain features (e.g., approximation coefficient), often lack a systematic and rigorous approach for selecting features that are directly and universally relevant to quality prediction and defect detection in laser welding. Additionally, the models designed for specific fault detection tasks have not been demonstrated to be adaptable and generalizable, which is critical when considering dynamic welding conditions and novel system configuration faults, such as unforeseen gaps between workpieces, variations in welding speed and laser power settings.

To enhance the explainability and generalizability of ML models in laser welding defect detection, this paper presents a rigorous framework for LWM data feature extraction and selection for ML model development, as depicted in **Fig.1**. The work aims at identifying the most valuable feature set to characterize defect-induced LWM signal variation, and developing an ML model that is generalizable for

dynamic welding conditions and evolving defect situations. The framework begins with the extraction of comprehensive features (covering time, frequency, and time-frequency domains) from LWM sensing signals, followed by correlation analysis and PCA for initial feature selection and ML model development. Upon the initial ML model-based defect classification, Shapley Additive Explanations (SHAP) is leveraged to perform a post-modeling analysis to identify features that significantly contribute to the model decision-making. Correspondingly, the initially selected feature set is further compressed, contributing to the complexity decrease and efficiency increase of the ML model. Subsequently, the adaptivity and generalizability of this ML model are enhanced through transfer learning, when the model is initially trained on data from a few defect categories but needs to be applied to detect more defect categories.

The rest of this paper is organized as follows: Section 2 details the feature extraction, selection, and ML classifier model, followed by Section 3 which presents the SHAP-based model explanation, future feature refinement, and improvement of the model's adaptivity employing transfer learning. Section 4 experimentally evaluates the performance of developed feature extraction and ML classification, and a conclusion is drawn in Section 5.

2. Feature-based ML model development

2.1 LWM sensing signal and feature extraction

In this study, the Precitec LWM system 4.0 is employed for inline welding process monitoring. The system's configuration is illustrated in **Fig.2 (A)**, which utilizes photodiode sensors to capture three types

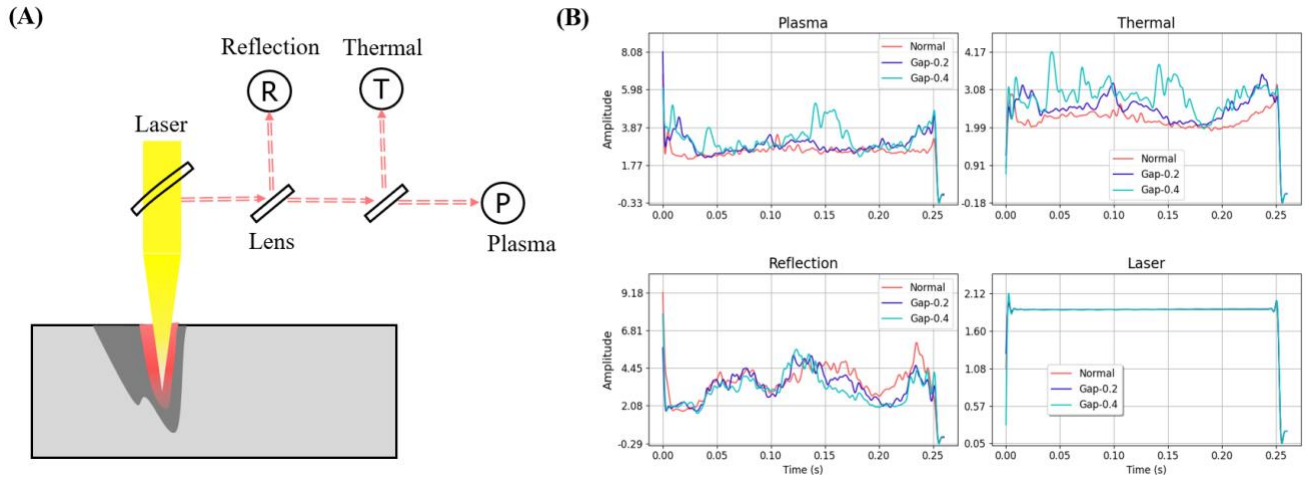


Fig.2 Sensing signals in LWM. (A) Setup of the laser welding monitoring system. (B) Time series sensing signals of plasma, thermal and back reflection, and comparison of these signals between normal welding and welding with abnormal gaps between material stackup.

of emission signals: plasma, thermal, and back reflection. The plasma channel captures ultraviolet light and luminous radiation from the plasma plume (450 nm to 580 nm), indicating the amount of metal vapor ionized during the keyhole formation process. The thermal signal samples the radiation in the near-infrared spectrum (1100 nm to 1800 nm), and the reflection captures unabsorbed laser radiation, matching the wavelength of the original laser beam (1020 nm to 1090 nm) (Weiss et al., 2022). Time-series sensing data have been recorded for normal weldings and various abnormal welding conditions, such as large gaps between material stack up, high and laser power, high and low welding speed and contaminated surface. An example of LWM signals collected from normal welding and welding with the abnormally large gap between material stack-up is shown in **Fig.2 (B)**. Then the goal of ML model development is to distinguish specific system configuration defects through the analysis of the sensing signals.

or each of the three LWM sensing channels (i.e., plasma, thermal, reflection), a comprehensive set of 23 physically meaningful features are first extracted in time, frequency, and time-frequency domains (Weeks, 2003), as summarized in Table 1. This leads to an initial feature pool with 69 features. Given the observed correlation between

the plasma and thermal signals considering the majority of these two channels' signal energy is contributed by the keyhole, weld pool and plume (Chianese et al., 2022), it is conceivable that some features might have considerable information overlapping. To remove redundant features, a Pearson's correlation analysis is performed to identify highly correlated feature groups (i.e., those with a correlation value exceeding 0.9), simplifying the feature pool from 69 to 32.

Table 1 Comprehensive feature pool

Time Domain	Frequency Domain	Time-Frequency Domain
Standard Deviation	Coefficient of Variability	Mean - Approximation Coefficient
Root Mean Square	Root Mean Square ratio	Std - Approximation Coefficient
Average	Average	Mean - Detail Coefficient 1
Skew	Mean Frequency	Mean - Detail Coefficient 2
Kurtosis	Skew	Std - Detail Coefficient 1
Crest Factor	Kurtosis	Std - Detail Coefficient 2
Latitude Factor	Stabilization Factor	Entropy
Shape Factor		
Impulse		

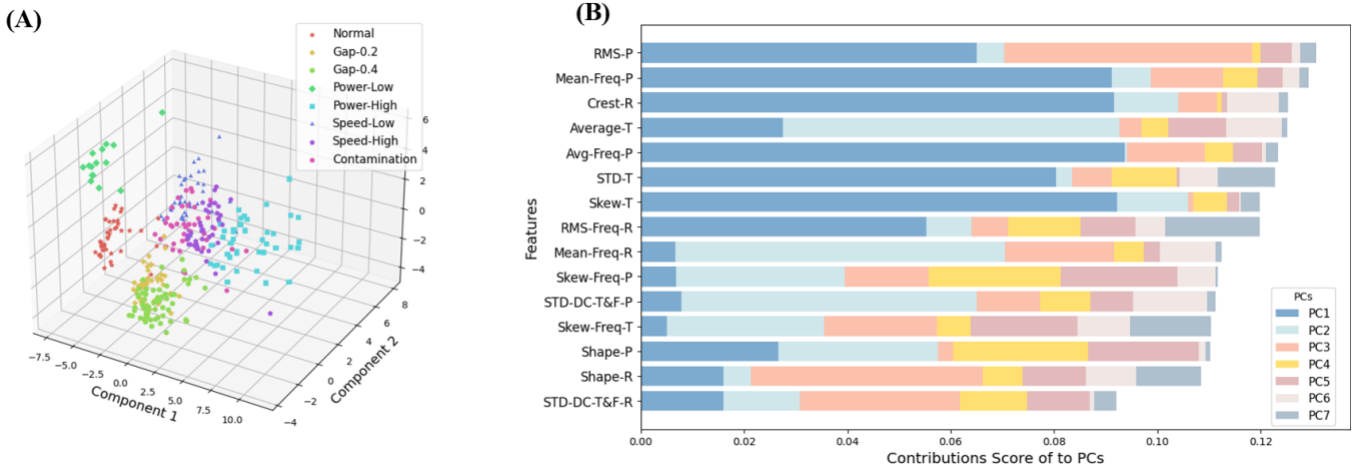


Fig.3 (A) Visualization of the PCs feature space. The samples from different defect categories are plotted with different colors. (B) The most important 15 features of PCs and the visualization of their contributions to each PC.

2.2 PCA-based feature selection

To further reduce the dimensionality of the feature pool, PCA is leveraged by identifying features that contributed to the construction of independent Principal Components (PC). Given the feature pool with 32 features selected from correlation analysis $F = [f_1, f_2, \dots, f_{32}]$ and normalized feature set that ensures uniformity across features $\hat{F}$, the PC space is constructed by solving:

$$\lambda_i b_i = \left(\frac{1}{n}\right) \hat{F}^T \hat{F} b_i \quad (1)$$

where λ_i represents the eigenvalue associated with the eigenvector b_i which is a column in the transformation matrix B . The matrix B transforms the normalized feature $\hat{F}$ into principal components by:

$$Z_i = b_i^T \hat{F}, \quad i = 1, 2, \dots, M \quad (2)$$

where Z_i are the calculated principal components (PCs) in the new feature space, and M is the total number of PCs derived.

Fig.3 (A) visualizes a 3-D PC space, where samples from eight system configuration defect categories are generally clustered together, demonstrating the effectiveness of the PCs. However, because of the unsupervised and data-driven essence of PCA, the PCs are not expandable to new datasets, as they would likely change with the appearance of new data (Peres-Neto et al., 2003).

To address this limitation and improve model robustness, it is essential to trace back the significance of the original features to the obtained PCs. By examining the loading scores in the PCA transformation matrix B , each feature's contribution to a PC can be determined. For a given PC PC_k , the contribution of an original

feature f_i can be expressed by the magnitude of its corresponding loading score B_{ik} . The contribution value C_{ik} can be represented by the absolute value of the loading score, which indicates the total variance in the original feature f_i is captured by the principal component PC_k . By summing these contributions across m PCs, the overall contribution $C_{i(total)}$ of an original feature f_i is calculated by:

$$C_{i(total)} = \sum_{k=1}^m C_{ik} \quad (3)$$

Based on this analysis, 15 features with the highest total contribution values are then selected, as presented in **Fig.3 (B)**, and an MLP model is then built using these features for defect detection.

2.3 MLP classifier modeling

Being the most fundamental neural network structure, MLP can approximate the nonlinear relationship between the inputs and outputs with the help of activation functions. Because of its structure simplicity, with limited capacity to directly process complex raw sensing data, an MLP's performance critically depends on the quality of the input features. This sensitivity to feature quality makes it a helpful tool for validating the effectiveness of the feature selection process. The architecture of the MLP, as depicted in **Fig.1 (C)**, begins with an input layer feeding with the 15 physical features, which are subsequently processed through two hidden layers, and finally outputs the probability of a sample belonging to one of the 8 defect categories.

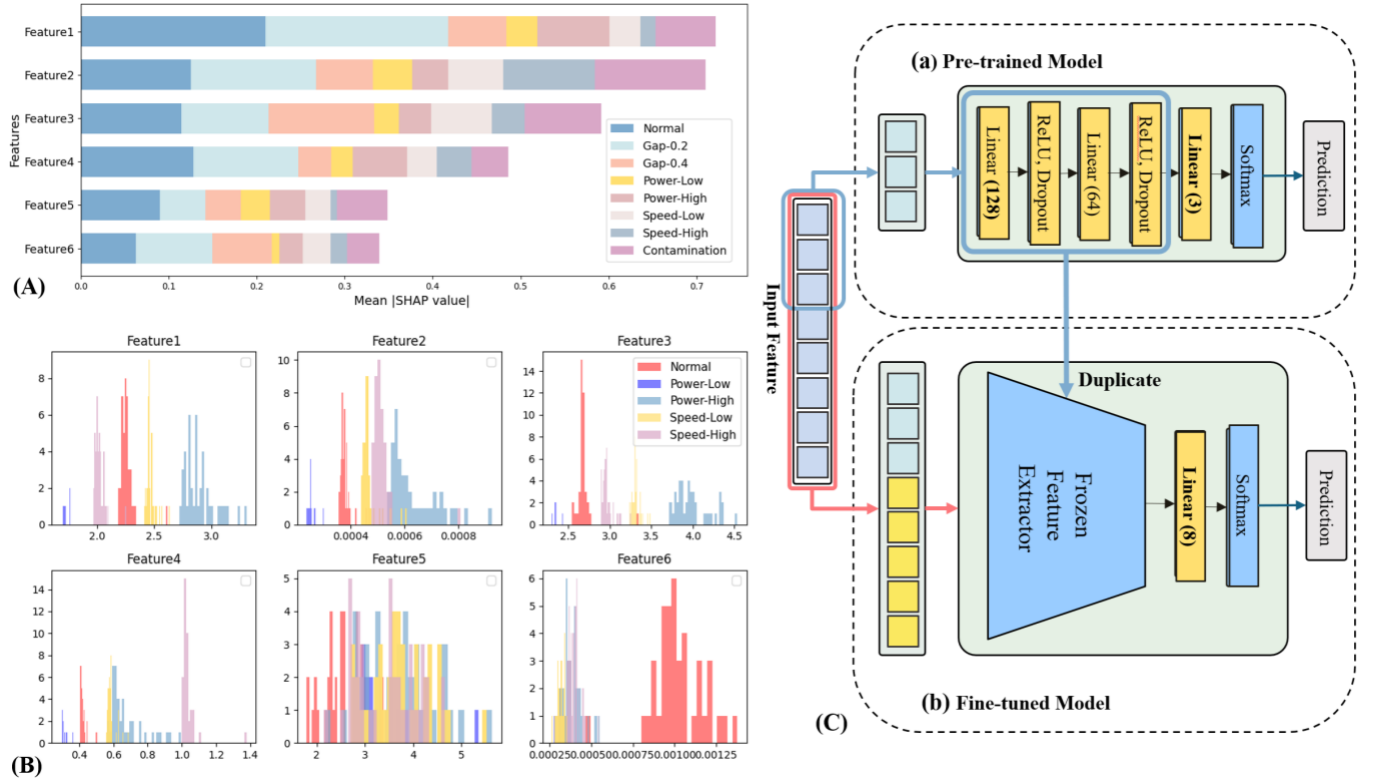


Fig.4 Analytical overview of model features and transfer learning framework. **(A)** Top six features with high SHAP value. Color variations in a single bar represent the contribution of a single feature to different conditions. **(B)** Distribution of the feature value under 5 conditions for six selected features. **(C)** The transfer learning framework in this study. The pre-trained **(a)** model is trained using data from three of the eight defect categories, and the fine-tuned model **(b)**, Model 3 listed in Table 2, is trained using all eight conditions to refine the output layer.

Experimental data including 290 weld samples in total are divided into two subsets: 203 samples for training and 87 samples for validation. This MLP model obtains a validation accuracy of 98.75%, as shown in Table 2 (Model 2), higher than the model training using the entire feature set of 69 features (Model 1), which is 96.25%.

3. Model explanation and improvement

3.1 Feature importance using SHAP analysis

While the MLP classifier using 15 extracted features has achieved high accuracy, the mechanism contributing to the success is unclear due to the black-box nature of the model. To improve the model's interpretability and further reduce the feature pool dimensionality, SHAP is implemented. Different from correlation and PCA analysis which mainly analyzes the inter-correlation among features, the SHAP analysis approaches the issue by quantifying the individual feature's contribution to the ML model output and decision-making.

SHAP calculates the marginal contribution, i.e., Shapley value, of each input feature to the output, inferred from the distinct output before and after a specific input's absence, expressed as:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|! (|N| - |S| - 1)!}{|N|!} (v(S \cup \{i\}) - v(S)) \quad (4)$$

where ϕ_i is the Shapley value for the feature i , N is the set of all features, which is the 15 input features in this case, and S is a subset of features excluding i . $|S|$ is the number of features in the subset S . $v(S)$ is the prediction function evaluated at the subset of features S . The sum is taken over all subsets S of the set $N \setminus \{i\}$. The standard method for calculating Shapley values, applied to linear prediction functions v , requires adaptations to suit nonlinear models such as the MLP used in this study. Therefore, an improved algorithm, Deep LIFT (Shrikumar et al., 2017), is leveraged to estimate the Shapley value by the gradients.

The SHAP value of a feature indicates its relative importance compared to others in the feature set to model outputs, thereby facilitating the selection of the influential features. Through SHAP analysis, 6 valuable features that contribute most to the model are identified, with their respective SHAP values illustrated in Fig.4 (A). In order to respect our industrial professional property interests, the specific names of these features have been obfuscated in this work.

A histogram describing the distribution of the feature value under five defect conditions is shown in Fig.4 (B). These selected features are closely correlated with specific physical characteristics of the process parameters. For example, one feature extracted from the thermal signal quantification effectively reveals variations in energy, reflecting the changes in process parameters, such as welding power and speed which influence the thermal dynamics during the welding process.

3.2 Transfer learning framework

Due to the complexities in the actual laser welding process, there will be varying operation and defect conditions while the training dataset is unable to cover all the system anomalies in data collection. Consequently, the adaptivity and generalizability of the model, i.e., the

ability to adapt from conditions seen during training to detect unseen conditions, is essential for its application in real-world welding settings. The adaptivity of this feature-based model depends on two factors. First, the input features must provide general insights to distinguish between various defect conditions, and second, the model must effectively extract the useful information from these features that can be applied to more welding scenarios.

Transfer learning facilitates the transfer of knowledge from one data domain to another, enhancing the model's ability to maintain performance across diverse application domains despite variances in data (Pan and Yang, 2010). In this context, transfer learning is applied to improve the generalizability of MLP built for defect detection, as illustrated in Fig.4 (C). Specifically, the MLP model initially pre-trained on a subset of the system defect categories, i.e., 3 of 8 categories, is expanded to recognize all 8 categories through transfer learning. To implement transfer learning, its knowledge learning block (the first two linear layers) is frozen, and only the last linear layer together with the softmax layer is fine-tuned with data from all 8 categories. Successful performance in this test, covering all conditions, would not only confirm the generalizability of the input features but also demonstrate the enhanced adaptability of the model.

4. Experiment and Result

The performance of the MLP models discussed in previous sections is evaluated on a consistent validation dataset, and the result is summarized in Table 2. Model 1, trained using all 69 features, achieves an accuracy of 96.25%, which is lower than that of the other models utilizing selected features. The diminished performance can be attributed to the presence of redundant or irrelevant features in the initial feature pool, which likely introduces confusion into the relatively simple MLP architecture. Model 4, trained with just 6 features which are identified through SHAP analysis of Model 2, maintains robust performance, validating the effectiveness of these selected features.

The confusion matrix for the validation of the transferred model (Model 3) and baseline (Model 4) is depicted in Fig.5. Model 3, which fine-tuned from a pre-trained model classifying 3 conditions with 99% accuracy, shows comparable prediction outcomes to Model 4, as the only miss-classified sample belongs to the same category of abnormal gap between material stackup, indicating that it is able to adapt from the training on a subset of conditions to identify unseen system configuration defects. Additionally, by reducing the number of inputs, Model 3 has only half the parameters of Model 1, which is 9,672 in total, significantly improving computational efficiency.

Table 2 Model performance

Model	Input Feature	Trained Condition	Predicted Condition	Model Size	Accuracy
Model 1	69	8	8	17736	96.25%
Model 2	15	8	8	10824	98.75%
Model 3	6	8	8	9672	98.75%
Model 4	6	3	8	9672	98.75%

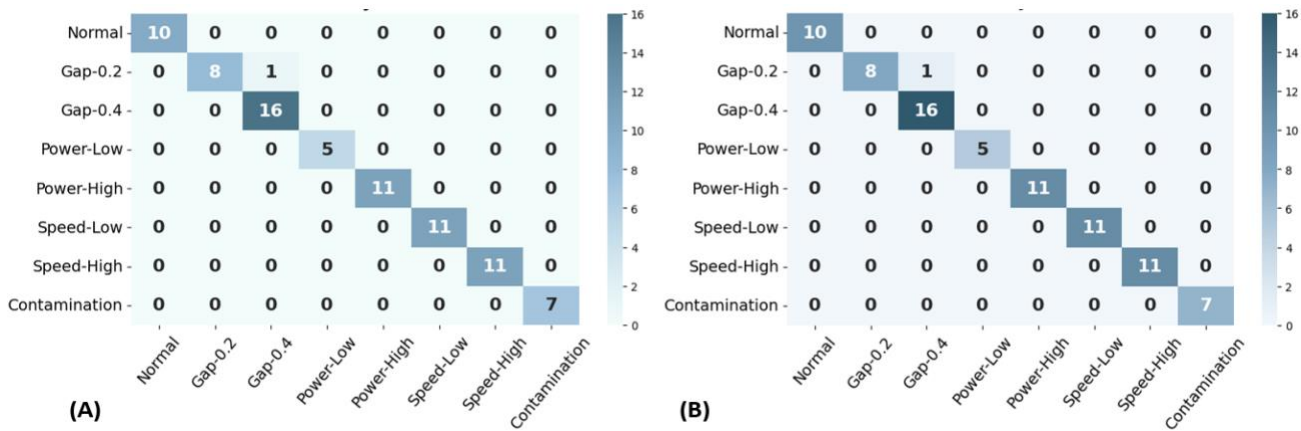


Fig. 5 Confusion matrices for the validation of the transferred Model 3 **(A)** and the baseline Model 4 **(B)**

5. Conclusion

This study proposes an adaptable machine learning framework utilizing a feature-based multi-layer perceptron for system configuration defect detection in laser welding. Through signal and post-modeling analysis, 6 physically meaningful features crucial for identifying system defects are determined and their effectiveness is demonstrated. Furthermore, transfer learning is utilized to improve the model's adaptability. The equivalence in the performance of the transferred model and the baseline model underscores its capacity to maintain model efficacy despite variations in training data and operational scenarios. Future work will expand the model to cover more system defects and dig deeper into the physical significance of the identified features, aiming to provide more profound insights into the mechanisms of laser welding.

ACKNOWLEDGEMENT

This paper is based upon work supported by the National Science Foundation under Grant No. 2237242.

REFERENCES

1. J. Feng et al. (2023) Application of Laser Welding in Electric Vehicle Battery Manufacturing: A Review. *MDPI* 13:1.
2. A. Sadeghian and N. Iqbal (2022) A review on dissimilar laser welding of steel-copper, steel-aluminum, aluminum-copper, and steel-nickel for electric vehicle battery manufacturing. *Optics and Laser Technology* 146:107595.
3. W. Cai et al. (2020) Application of sensing techniques and artificial intelligence-based methods to laser welding real-time monitoring: A critical review of recent literature. *J. Manuf. Sys.* 57: 1-18.
4. E. A. I. Eriksson et al. (2009) Basic study of photodiode signals from laser welding emissions. *Nordic Conference on Laser Processing of Materials*.
5. C. Angeloni et al. (2024), Laser welding in e-mobility: process characterization and monitoring. *Manuf. Materials Proc.* 11:1:3-24.
6. J. Li et al. (2022) Prediction of penetration based on plasma plume and spectrum characteristics in laser welding. *J Manuf Process* 75:539-604.

7. Y. Zhang et al.(2019) Welding defects detection based on deep learning with multiple optical sensors during disk laser welding of thick plates. *J Manuf Syst* 51:87-94.
8. K. She et al.(2024) Online Detection of Laser Welding Penetration Depth Based on Multi-Sensor. *Materials* 17:7.
9. T. Brežan et al. (2023)Fusing optical coherence tomography and photodiodes for diagnosis of weld features during remote laser welding of copper-to-aluminum. *J Laser Appl* 35:1.
10. H. Kim et al. (2022) Analysis of laser-beam absorptance and keyhole behavior during laser keyhole welding of aluminum alloy using a deep-learning. *J Manuf Process* 80:75-86.
11. T. Will et al. (2022) Feature extraction based on scalable hypothesis tests from photodiode data in laser welding processes. *Procedia CIRP* 527–531.
12. P. Stavropoulos et al (2022) Quality assurance of battery laser welding: A data-driven approach. *Procedia CIRP* 784–789.
13. G. Chianese et al. (2022) Using photodiodes and supervised ML for automatic classification of weld defects in laser welding of thin foils copper-to-steel battery tabs. *J Laser Appl* 34:4.
14. T. Weiss et al. (2022) A holistic approach for an intelligent laser beam welding architecture using machine learning for the welding of metallic bipolar plates for polymer electrolyte membrane fuel cells. *Procedia CIRP* 810–815
15. M. Weeks (2003) Discrete Wavelet Transform: Architectures. Design and Performance Issues. *J Signal Processing* 35: 155-178.
16. G. Chianese et al. (2022) Characterization of Photodiodes for Detection of Variations in Part-to-Part Gap and Weld Penetration Depth During Remote Laser Welding of Copper-to-Steel Battery Tab Connectors. *J of Manuf. Sci. Engineering* 144:7.
17. P. R. Peres-Neto et al. (2003) Giving meaningful interpretation to ordination axes: Assessing loading significance in principal component analysis. *Ecology* 84: 9: 2347–2363.
18. Shrikumar et al. (2017) Learning important features through propagating activation differences. In *Proceedings of the 34th international conference on machine learning* 70: 3145–3153.
19. S. J. Pan and Q. Yang (2010) A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering* 22: 1345–1359.