Applications

Patricia M. Dold*, Fabian Bleier, Meiko Boley and Ralf Mikut

Two-stage quality monitoring of a laser welding process using machine learning

Zweistufige Qualitätsüberwachung eines Laser-Schweißprozesses mit Hilfe maschinellen Lernens

An approach for fast yet precise quality monitoring

Abstract: In production, quality monitoring is essential to detect defective elements. State-of-the-art approaches are single-sensor systems (SSS) and multi-sensor systems (MSS). Yet, these approaches might not be suitable: Nowadays, one component may comprise several hundred meters of the weld seam, necessitating high-speed welding to produce enough components. To detect as many defects as possible in time, fast yet precise monitoring is required. However, information captured by SSS might not be sufficient and MSS suffer from long inference times. Therefore, we present a confidence-based cascaded system (CS). The key idea of the CS is that not all data are analyzed to obtain the quality weld, but only selected ones. As evidenced by our results, all CS outperform SSS in terms of accuracy and inference time. Further, compared to MSS, the CS has hardware advantages.

Keywords: two-stage quality monitoring; cascaded system; single- and multi-sensor system; decision trees; neural networks; intelligent process monitoring

Zusammenfassung: Qualitätsüberwachung ist in der Produktion unerlässlich, um fehlerhafte Elemente zu erkennen. Zum Stand der Technik gehören Single-Sensor-Systeme (SSS) und Multi-Sensor-Systeme (MSS). Diese Ansätze sind jedoch möglicherweise nicht geeignet: Heutzutage kann ein Bauteil mehrere hundert Schweißnaht-Meter

umfassen, weshalb hohe Schweißgeschwindigkeiten erforderlich sind, um genügend Bauteile zu produzieren. Um möglichst viele Fehler rechtzeitig zu erkennen, ist eine schnelle und dennoch präzise überwachung erforderlich. Die von SSS erfassten Informationen sind möglicherweise nicht ausreichend und MSS leiden unter langen Inferenzzeiten. Daher stellen wir ein vertrauensbasiertes kaskadiertes System (CS) vor. Der Grundgedanke des CS ist, dass nicht alle Daten analysiert werden, um die Schweißnahtqualität zu erhalten, sondern nur ausgewählte Daten. Unsere Ergebnisse zeigen, dass alle CS das SSS in Bezug auf Genauigkeit und Inferenzzeit übertreffen. Außerdem hat das CS im Vergleich zum MSS Hardware-Vorteile.

Schlagwörter: zweistufige Qualitätsüberwachung; kaskadiertes System; Single- und Multi-Sensor-System; Entscheidungsbäume; neuronale Netze; intelligente Prozessüberwachung

1 Introduction

Laser welding is used in automotive [1, 2], aerospace [3], or shipbuilding [4] industries and is considered a key production technology [5–7]. Its ability for precise and fast welding is advantageous compared to other joining techniques. During a laser welding process, material is melted, vaporized and solidified. Unfortunately, this process often is challenging [8] and, hence, welding defects like false friend [9], humping [10], undercut [11] or dropout occur. Especially in the electric drive train, some components consist of several hundred laser-welded elements, and one welding defect can lead to the failure of the entire component. To quickly detect defective elements, quality monitoring is desired in industrial processes.

Two main sensors to evaluate the weld quality are photodiodes (PD) [12–15] and spectrometers [16, 17] because of their simple structure. Other sensors are ultraviolet sensors

*Corresponding author: Patricia M. Dold, Bosch Research, Robert Bosch GmbH, Robert-Bosch-Campus 1, 71272 Renningen, Germany; and Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany, E-mail: patricia.dold@de.bosch.com
Fabian Bleier and Meiko Boley, Bosch Research, Robert Bosch GmbH, Robert-Bosch-Campus 1, 71272 Renningen, Germany
Ralf Mikut, Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

[18], X-ray sensors [19, 20], microscopy, microphones [19], optical coherence tomography [21], and high-speed cameras (HSC) [22–25]. Accordingly, compared to typical quality monitoring in resistance or ultrasonic welding [26, 27], data from more sensors exist.

The signals acquired by the different sensors are analyzed with processing algorithms. Thereby, data-driven process monitoring has been implemented by applying machine learning methods like support vector machines [28–30], decision trees (DT) [31], random forest algorithms [32, 33], or Bayes classifiers [34]. Recently, deep learning has achieved success in image recognition and classification [35] and thus has been applied to weld defect inspection [25, 36, 37].

Some researchers use a single sensor type, or single sensor system (SSS), to study a mechanism of the welding process. In [13], the optical intensity captured by a PD when welding defects occur is analyzed. In [25], the results of quality assessment of HSC or optical coherence tomography data with neural networks (NN) like Inception-v3 [38] are presented. Moreover, different convolutional neural networks (CNN) like AlexNet [35], VGG-16 [39], or MobileNetV3-Large [40] are used for the automated optical inspection of a laser welding process [36].

However, information captured by one sensor is not sufficient for holistic quality assessment [41]. A combination of multiple sensor types, on the other hand, provides a more comprehensive description of the welding process [19]. Consequently, multi-sensor systems (MSS) use different sensors [19, 29, 30, 33, 37, 42]. For example, during high-power disc laser welding, a PD, a spectrometer, and an ultraviolet sensor can be used to detect different weld defects [43].

A quality monitoring system whose inference time is not longer than its cyclic time is desired because the production process is not slowed down. Yet, because of long inference times, complex MSS and processing algorithms can result in quality monitoring approaches that are not profitable in production. In other domains, similar problems were solved by two-stage classifiers, like for solder joint inspection [34, 44], surface-mounted devices inspection [45], or ball bonding [46].

The present paper extends our conference paper [47], which presented a cascaded system (CS) with the aim of fast but still precise quality monitoring. The CS follows a two-stage structure: The first level of inspection analyzes simple data like time series, with the advantage of a high clock rate and low memory requirements. This stage already safely classifies some welds; in areas of uncertainty, however, the next more complex stage, which may use image

data, takes over to make a final decision. The extensions include:

– usage of deep NN architectures like CNN or a multilayer perceptron (MLP) for PD time series in the first stage of the CS and Inception-v3 [38] or ResNet50 [48] for HSC images in the second stage,
– usage of classical machine learning algorithms including feature engineering and DT with the advantage of more interpretable results,
– comparison of 16 NN-based CS and 4 DT-based CS, so in total 20 different CS as well as their corresponding SSS and MSS, and
– Pareto optimization with respect to the accuracy and the inference time for quality monitoring approach design at given conditions.

2 Data set

The data were acquired in the laboratory. Photodiode (PD) signals and synchronous high-speed camera (HSC) images were captured during the welding of thin metal plates. Subsequently, the data were preprocessed so that they can be used by the quality monitoring approaches described in Section 3.

2.1 Experimental setup

Figure 1 shows the experimental setup. Thereby, a laser beam generated in the laser was directed onto two thin metal plates by a 2D galvanometer scanner using two mirrors. As a laser, an IPG YLR-1000-WC-Y14 fiber laser with an infrared wavelength of 1070 nm at a power of 250 W was used. As a scanner, a Scanlab intelliWELD PR, which focused the beam to a diameter of 45 μm on the workpiece surface, was used. Two measurement systems observed the
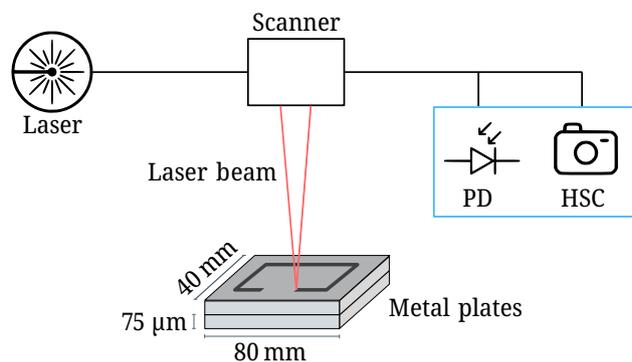
Figure 1: Experimental setup of the laser welding process of two thin metal plates. Two measurement systems were used: a photodiode (PD) and a high-speed camera (HSC).

laser welding process coaxially and in-process: Firstly, a PD (FEMTO LCA-S-400K) measured light with a wavelength of 300–950 nm in the area of the weld pool. Accordingly, at each sampling time, the voltage of the PD amplifier was recorded. Secondly, HSC (Optronics Cyclone-1HS-3500) images were recorded.

2.2 Experiments

As shown in Figure 1, two identical metal plates with a length of 80 mm, a width of 40 mm and a thickness of 75 µm were welded together in an overlap joint. The plates were clamped and full penetration welded in a rectangular geometry. After welding, the two plates should be joined along the rectangular path. The welding process for one metal plate pair took 338 ms. In total, 59 pairs were welded. Thereby, 9 were welded under reference conditions, whereas anomalies like spatters or gaps were inserted in 50.

Figure 2 shows the data and the sampling rates of the two observing sensors. Thereby, the axis indicates the time in µs. Above the axis, the PD time series is plotted. The PD had a sampling rate of 250 kHz, so every 4 µs one sample was recorded. Below the time axis, the HSC images are shown. The HSC had a sampling rate of 20 kHz, resulting in one image every 50 µs. The higher sampling rate of the PD allows the detection of shorter anomalies. Moreover, faster processing is possible because of the smaller amount of raw data compared to the HSC images. However, the HSC images provide information that is not in the PD signals, such as information about the interaction zone geometry

because the PD signals are spatially aggregated signals, whereas the HSC images have a spatial resolution. Moreover, Figure 2 shows reference and anomaly data. On the left side (green box), the process was welded under reference conditions. On the right side (red box), oil was inserted between the metal plates, leading to an anomaly. Inserting the oil resulted in a difference in the amplitude in the PD time series and in a slight difference in the interaction zone geometry in the HSC images.

2.3 Data preprocessing

The grayscale HSC images were cropped to a size of 100×100 pixels and scaled to a value range of [0,1]. Then, each HSC image was assigned to 13 PD samples, representing a chunk. These chunks resulted from the sampling rates of the sensors and are visualized in Figure 2 by the gray boxes. Next, to each chunk a label, namely reference or anomaly, was assigned. By such chunk-wise labeling, defect localization along the welding path is possible, because different locations on one metal plate pair have different labels. An anomaly refers to those locations, where an abnormality like a gap or spatters was introduced. Reference refers to the locations, where neither an abnormality was provoked nor visible in the recorded PD signals or HSC images. Accordingly, Table 1 shows the number of chunks of the data set. For the evaluation of the models, 5-fold cross-validation was applied. Thereby, the training and test splits are based on whole metal plate pairs. Such a split is closer to the production scenario as algorithms are trained on data of some workpieces and then applied to data of other ones. However, a random split would lead to a more similar distribution of the training and test data. Moreover, to have metal plate pairs of each error case in each fold, a division of the data into 5 folds is reasonable. The average number of training chunks over the 5 folds is given by n_{train}, the average number of test chunks by n_{test}, and both together as n. To teach the models invariances and robustness properties, a data augmentation on the HSC images including rotation and to flipping was used. Accordingly, the average chunk number with data augmentation is given by n_{DA}.
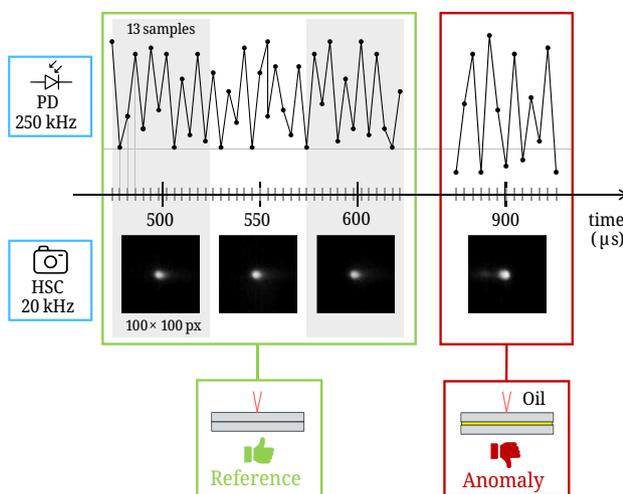
Figure 2: Data and sampling rates of the PD and the HSC. Thereby, reference and anomaly data are indicated by the green and red boxes; the assignment of the PD and HSC data as chunks is given by the gray boxes.

Table 1: Number of chunks of the data set. Given are the average numbers over the 5 folds of the cross-validation.

Label	n_{train}	n_{test}	n	n_{DA}
Reference	148 425	35 760	184 185	1 605 716
Anomaly	147 368	38 188	185 556	1 594 284
Total	295 793	73 948	369 741	3 200 000

3 Quality monitoring approaches

Three different quality monitoring approaches, namely single-sensor system (SSS), multi-sensor system (MSS), and cascaded system (CS) are considered. In general, the approaches map measurement values X_i to a quality-relevant quantity Y_i. Before the approaches are defined, the formal definitions of the data are given:

Each out of n PD time series is defined as

$$X_{\mathrm{PD},i} = (x_{i,1} \quad x_{i,2} \quad \ldots \quad x_{i,13}),$$

where the first index $i \in \{1, \ldots, n\}$ indicates the chunk number and the second index the sample number within that chunk. The images of the HSC are defined as

$$X_{\mathrm{HSC},i} = \begin{pmatrix} x_{i,1,1} & x_{i,1,2} & \ldots & x_{i,1,100} \\ x_{i,2,1} & x_{i,2,2} & \ldots & x_{i,2,100} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i,100,1} & x_{i,100,2} & \ldots & x_{i,100,100} \end{pmatrix}.$$

With the label $Y_i \in \{0, 1\}$, where 0 indicates anomaly and 1 reference, the data set S consists of $n = 369\,741$ (see Table 1) triples according to

$$S = \{(X_{\mathrm{PD},1}, X_{\mathrm{HSC},1}, Y_1), \ldots, (X_{\mathrm{PD},n}, X_{\mathrm{HSC},n}, Y_n)\}.$$

3.1 Single-sensor system

A SSS performs process monitoring based on data coming from one sensor. For the PD and HSC data as input, the SSS are defined by the functions

$$f_{\mathrm{SSS,PD}}\colon \mathbb{R}^{1\times13} \to \{0,1\}\colon X_{\mathrm{PD},i} \mapsto Y_i,$$

$$f_{\mathrm{SSS,HSC}}\colon \mathbb{R}^{100\times100} \to \{0,1\}\colon X_{\mathrm{HSC},i} \mapsto Y_i.$$

Figure 3 shows the two SSS: The one with PD signals as input (left side) and the one with HSC images as input (right side).
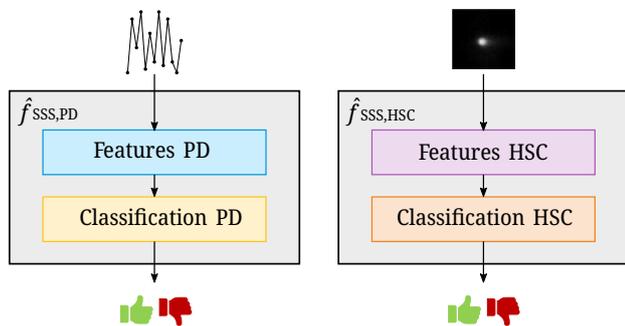
Thereby, $\hat{f}_{\mathrm{SSS,PD}}$ and $\hat{f}_{\mathrm{SSS,HSC}}$ indicate optimized prediction models. Each prediction model consists of two blocks: A features block to extract the most important features and a classification block to determine the weld quality. Concrete algorithms for each block, namely DT and NN, are given in Section 4.

3.2 Multi-sensor system

A MSS uses data from multiple sensors. A MSS consisting of PD and HSC data as input is expressed by

$$f_{\mathrm{MSS}}\colon \mathbb{R}^{1\times13} \times \mathbb{R}^{100\times100} \to \{0,1\}\colon X_{\mathrm{PD},i} \times X_{\mathrm{HSC},i} \mapsto Y_i.$$

Figure 4 shows a MSS with its prediction model $\hat{f}_{\mathrm{MSS}}$. The MSS consists of two feature blocks, namely one for each sensor data, and one classification block, where the features are fused and processed for prediction. Thereby, the classification block has the algorithm structure like the one of the SSS,HSC. However, the difference is that the input dimension changes because PD features are also processed.

3.3 Cascaded system

The advantage of MSS compared with SSS is a more holistic quality assessment, however, the inference time of MSS is greater. Therefore, the CS combines the advantages of both systems. It offers the possibility to use multiple sensors as a MSS. In contrast to MSS, not all data are analyzed to obtain the weld quality; only selected ones. Figure 5 shows a two-stage CS. Let $p \in [0,1]$ be the output of a classifier that predicts based on PD signals. For $p < 0.5$ the classifier chooses anomaly; for $p \geq 0.5$ reference. The closer p is to 0 or 1, the more confident the classifier's decision is considered. Let $r \in (0, 0.5)$ be a fixed threshold. A classifier's decision is certain if $p < r$ or $p > 1 - r$. If the first condition is satisfied, the classifier is certain for anomaly; if the second is satisfied,

Figure 3: Two single-sensor systems (SSS). Left side: SSS with $X_{\mathrm{PD},i}$ as input and $\hat{f}_{\mathrm{SSS,PD}}$ as optimized prediction model. Right side: SSS with $X_{\mathrm{HSC},i}$ as input and the optimized prediction model $\hat{f}_{\mathrm{SSS,HSC}}$.
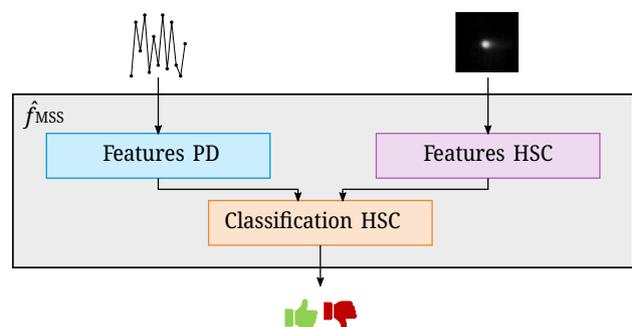
Figure 4: Multi-sensor system (MSS). It has $X_{\mathrm{PD},i}$ and $X_{\mathrm{HSC},i}$ as input and $\hat{f}_{\mathrm{MSS}}$ as prediction model.
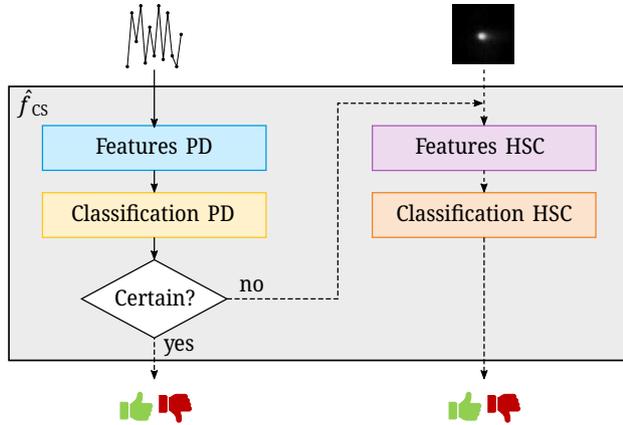
Figure 5: Cascaded system (CS). Depending on the certainty, which results from p and r, it uses only $X_{PD,i}$, or $X_{PD,i}$ and $X_{HSC,i}$ as input. The prediction model is $\hat{f}_{CS}$.

it is certain for reference. If the classifier is certain, its result is accepted; if not, a final decision is made in a next step based on the HSC data. Formally, the CS is expressed by

$$f_{CS} : \begin{cases} f_{SSS,PD}, & \text{if} \quad p < r \text{ or } p > 1 - r \\ f_{SSS,HSC}, & \text{otherwise.} \end{cases}$$

4 Implementation details

The prediction models $\hat{f}_{SSS,PD}$, $\hat{f}_{SSS,HSC}$, $\hat{f}_{MSS}$, and $\hat{f}_{CS}$ introduced in Section 3 were created by combining four blocks: Features PD, Features HSC, Classification PD, and Classification HSC (see Figures 3–5). Building the approaches with same blocks ensures a fair comparison. For each of the four blocks, different algorithms were implemented. Thereby, on one hand, blocks consist of classical machine learning methods, namely feature engineering and decision trees (DT), and, on the other hand, of neural networks (NN). The algorithm structure of the blocks are given in Tables 2 and 3. Thereby, one feature and classification block form a configuration. In total, two DT- and four NN-based configurations for each SSS,PD and SSS,HSC were considered. The MSS and CS result from combining each DT-configuration of Table 2 with each DT-configuration of Table 3 resulting in $2^2 = 4$ different DT-based MSS and CS. Analogously, each NN-configuration of Table 2 was combined with each NN-configuration of Table 3, resulting in $2^4 = 16$ different NN-based MSS and CS. Consequently, in total, 20 MSS and 20 CS were considered. The concrete algorithms of each configuration are explained in the follow-

ing. Thereby, hyperparameter optimization, namely a grid search, and feature selection was conducted on the training data.

4.1 Feature engineering and decision trees

In the following, configurations DT1–DT4 from Tables 2 and 3 are explained. Features from the PD signals were extracted in two different ways: manually and automated with the python toolbox tsfresh [49]. For the manual feature extraction seven statistical (stat.) features were calculated on every of the 13 samples in a chunk of the PD data. The seven features were: mean, standard deviation, maximum, minimum, distance between maximum and minimum, kurtosis, and skewness. To not have overlooked any important information, also an automated feature extraction using the python toolbox tsfresh was used. The toolbox automatically calculates 794 time series features.

From each HSC image, the following statistical and geometrical (geo.) features were manually selected: Firstly, the seven statistical features as calculated on the PD signal

Table 2: Configurations processing PD signals. Each configuration consists of a Features PD and a Classification PD block. The DT-based configurations are explained in Section 4.1 and the NN-based in Section 4.2.

Configuration	Features PD	Classification PD
DT1	Manual stat.	$d = 8, s = 200$
DT2	tsfresh [49]	$d = 8, s = 200$
CNN1	$C_1 2^6$-$C_1 2^7$-$C_1 2^8$-$C_1 2^6$	F-D2^6-P-D2^4-P-D2^3-P-D1
CNN2 [47]	$C_1 2^3$-$C_1 2^4$-$C_1 2^4$-$C_1 2^3$	F-D2^4-P-D2^3-P-D2^2-P-D1
CNN3	$C_1 5$-$C_1 5$-$C_1 5$	F-D1
MLP	–	D5-D5-D5-D1

Table 3: Configurations processing HSC images. Each configuration consists of a Features HSC and a Classification HSC block. The DT-based configurations are explained in Section 4.1 and the NN-based in Section 4.2.

Configuration	Features HSC	Classification HSC
DT3	Manual stat.+geo.	$d = 16, s = 200$
DT4	Manual stat.+geo.	$d = 32, s = 2000$
MN [47]	MobileNet [50][a]	F-D2^8-P-D2^8-P-D2^7-P-D1
IV3	Inception-v3 [38][a]	F-D2^8-P-D2^8-P-D2^7-P-D1
RN50	ResNet50 [48][a]	F-D2^8-P-D2^8-P-D2^7-P-D1
CNN4	$C_2 2^5$-$C_2 2^6$-$C_2 2^7$-$C_2 2^6$	F-D2^6-P-D2^5-P-D2^4-P-D1

[a]Only part before dense layers.

chunks were calculated for each HSC image as well as the median. Secondly, depending on a threshold $h \in [0, 255]$, eleven geometrical features were calculated. Therefore, a binary mask image with the same dimension as the HSC images was calculated for every HSC image. The mask has the value 1 at those locations where the pixel values in the HSC image are greater than h or equal, and 0 where the HSC image pixel values are smaller. Depending on h, the binary mask has information about for example the keyhole size, the keyhole shape, or spatters. Based on the binary mask, the following eleven features were extracted: area, number of regions, area of the biggest region, ratio of area of the biggest region and area, convex hull, ratio of area and convex hull, circumference, ratio of circumference and area, area of a fitted ellipse, length of the ellipse and width of the ellipse. Fifteen different empirical thresholds h were considered. This results in $11 \cdot 15 = 165$ geometrical features per image. Together with the eight statistical features, a total of 173 features of each HSC image were considered.

After feature extraction, classification was done with DT. The DT were implemented with the library scikit-learn [51]. The library uses an optimized version of the CART algorithm. To measure the quality of a split, the Gini impurity was used as criterion. A grid search to find good hyperparameters, namely the maximum depth of the tree d and the minimum number of chunks required to split an internal node s, was performed. The specific values of the hyperparameters d and s can be taken from Tables 2 and 3. To obtain the most relevant features, the importance of each feature based on the Gini importance was calculated. Reducing to the most important features highly decreases the time to build the tree. Only those features were selected to build the final DT that were at least under the ten most relevant features in one of the trees of the 5-fold cross-validation.

4.2 Neural networks

In the following, all other configurations from Tables 2 and 3, namely CNN1–CNN4, MLP, MN, IV3, and RN50, are described. All consist of different NN. Thereby, the blocks Features PD and Features HSC consist of convolution layers whereas the blocks Classification PD and HSC consist of dense layers. The concrete architectures are given in Tables 2 and 3.

The Feature PD blocks of CNN1–CNN3 consist of $C_1 k$. $C_1 k$ describes a 1D-convolution layer with a filter size of 1×3 followed by batch normalization, and a ReLU as activation function. $k \in \mathbb{N}$ indicates the number of filters. 1D-convolutions were used on the PD time series because the input is 1-dimensional. CNN4 consists of $C_2 k$. $C_2 k$ describes

a 2D-convolution layer with a filter size of 3×3 followed by batch normalization, max pooling, and a ReLU as activation function. MN, IV3, and RN50 use the convolution part of the existing architectures MobileNet [50], Inception-v3 [38], and ResNet50 [48], respectively.

Regarding the classification blocks, F describes a flatten layer. Dl describes a dense layer with $l \in \mathbb{N}$ neurons. Thereby, a ReLU activation function is used if $l \neq 1$ and a Sigmoid activation function if $l = 1$. P describes a dropout layer with a dropout rate of 0.5.

During training of the NN, the binary cross-entropy was used as loss function. An Adam optimizer with a learning rate $l_r = 5 \cdot 10^{-5}$ and the regularizers $\beta_1 = 0.9$ and $\beta_2 = 0.99$ was used. The NN were trained with a batch size of 32 with 1000 steps per epoch for 100 epochs and with the same seed for better performance comparison. All NN were initialized with random weights. Additionally, 5-fold cross-validation was used to obtain the final results. The NN were implemented in Python using Keras [52] and Tensorflow [53]. The training processes ran on an NVIDIA A40 GPU card. In total 120 NN were trained (taking into account the 5-fold cross-validation so 24 NN in each fold): 20 $\hat{f}_{SSS,PD}$ with a training time of 3–6 min each, 20 $\hat{f}_{SSS,HSC}$ and 80 $\hat{f}_{MSS}$ with a training time of 2–4 h each.

4.3 Certainty of the cascaded system

Besides the four blocks Features PD, Features HSC, Classification PD and Classification PD, the CS consists of a request if the PD classification result is certain or not (see Figure 5). For the 16 CS consisting of NN, the output of the NN before binarization is used. This output is a value between 0 and 1 because of the binary classification. Model calibration is not necessary, as a certainty ranking of the chunks is sufficient. However, if interpretable r and p values as probabilities are needed, model calibration like plat scaling or isotonic regression would be necessary. For the 4 CS consisting of DT-based methods, the certainty is given by the fraction of the chunk belonging to the same class in a leaf. For example, if all chunks in a leaf belong to anomaly or reference, the certainty is considered as 0 or 1, respectively. If half of the chunks in a leaf belong to reference and the other half to anomaly, then the certainty value for the chunks in that leaf would be 0.5.

4.4 Inference time

The quality monitoring approaches are compared with respect to the accuracy and the inference time. In [47] the number of parameters is used to estimate the inference time. However, only NN have been considered. The present

paper deals with different types of algorithms, namely feature calculations, DT and NN, so that parameter comparison is not feasible. Therefore, the inference time on the same hardware, a CPU (AMD EPYC 7543 32-core processor), is used. Thereby, the mean inference time per chunk of the test chunks was calculated. The mean evaluation time is used for the following reason: In production, the cyclic time determines the evaluation time to remain real-time capable. If the evaluation time is limited to the cyclic time for each element, there probably are elements with shorter evaluation times and elements with longer evaluation because of the cascaded structure of the CS. A shorter evaluation means that there is time when no evaluation takes place. Longer evaluation needs to be interrupted when the cyclic time is reached. In contrast, if evaluation takes place with an offset time t_o after the welding process, then the mean evaluation time per element over the time t_o must be shorter than the cyclic time. The advantage is that evaluation times shorter than the cyclic time of one element allow longer evaluation times for other elements. It must be mentioned that the implementation of the algorithms is not optimized yet, and the inference times are considered as a rough indication.

5 Results and discussion

Figure 6 shows the results of the different approaches. Thereby, the accuracy is shown over the inference time. The accuracy is given in % on a linear scale, while the inference time is given in ms on a logarithmic scale. In the graph, yellow symbols represent SSS,PD, red symbols SSS,HSC, and blue symbols MSS. Gray lines indicate the performances of CS depending on the threshold r. Thereby, the threshold r was changed in steps of 0.005. Each black symbol lies on one of the lines and indicates the CS with the highest accuracy. While the color of a symbol characterizes the approach, the symbol's shape stands for the configuration (see Tables 2 and 3). The MSS and CS symbols (blue and black) consist of two parts: an inner and outer symbol. The inner symbol indicates the Features PD and Classification PD block. The outer symbol indicates the Features HSC and Classification HSC blocks (see Figures 4 and 5). As shown in Figure 4, the MSS does not consist of a Classification PD block. Therefore, for MSS, the inner symbol refers only to the Features PD block. The exact values of the accuracy and the inference time are given in Table 4. Additionally, Table 4 includes precision, recall, and F1-score with standard deviations.
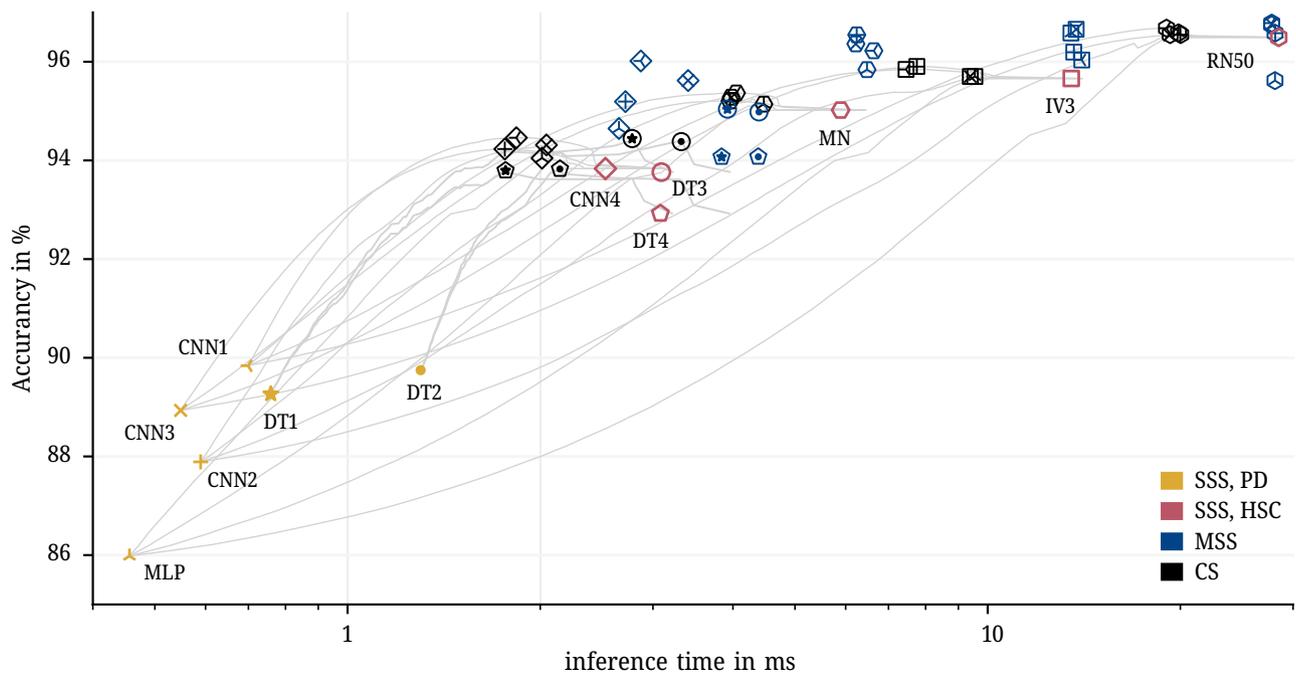
Figure 6: Performance of the approaches and their configurations with respect to the accuracy and the inference time. The four approaches are marked by the four colors: yellow, red, blue and black. The configurations are marked by the symbols. The configurations of SSS,PD and SSS,HSC (see Tables 2 and 3) are written next to the corresponding symbol. The symbols of the MSS and the CS result from combining the SSS-configurations. All symbols with their corresponding configurations can also be seen in Table 4. The gray lines indicate the performance of the CS depending on the threshold r, where the CS-symbols indicate the maximum value on each line.

Table 4: Results of each configuration of each approach with 5-fold cross-validation including accuracy, precision, recall, F1-score, and inference time. Thereby, of each approach, the maximal accuracy, precision, recall, and F1-score as well as the minimal inference time are marked in bold. The best DT-configurations of each approach are marked in gray bold and the best NN-configurations in **black bold**. For the accuracy, precision, recall, and F1-score the standard deviation is also given. The entries of the CS result from where the accuracy of each CS-configuration has its maximum.

Approach	Configuration	Symbol	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Inference time (ms)
SSS,PD	DT1	★	89.27 ± 2.12	85.79 ± 4.60	94.19 ± 0.68	89.72 ± 2.33	**0.759**
	DT2	•	**89.75 ± 1.74**	**86.33 ± 3.92**	**94.80 ± 0.54**	**90.17 ± 2.00**	1.301
	CNN1	⊰	**89.84 ± 1.92**	**86.86 ± 4.18**	94.02 ± 2.11	**90.21 ± 2.06**	0.700
	CNN2 [42]	✕	87.89 ± 2.73	84.01 ± 5.97	94.34 ± 3.91	88.63 ± 2.50	0.590
	CNN3	+	88.93 ± 1.82	85.38 ± 3.45	93.90 ± 1.66	89.40 ± 2.16	0.549
	MLP	⊥	85.98 ± 2.42	80.56 ± 4.01	**94.92 ± 0.83**	87.12 ± 2.49	**0.456**
SSS,HSC	DT3	⬡	**93.76 ± 1.79**	**96.02 ± 1.65**	**94.42 ± 1.90**	**95.19 ± 0.74**	3.092
	DT4	⬠	92.92 ± 1.85	94.97 ± 1.39	93.52 ± 1.87	94.51 ± 0.89	**3.081**
	MN [42]	⬡	95.02 ± 2.33	93.52 ± 4.00	97.05 ± 5.72	95.21 ± 2.15	5.893
	IV3	☐	95.66 ± 2.37	**96.46 ± 0.93**	94.92 ± 1.16	95.68 ± 2.70	13.509
	RN50	⬡	**96.50 ± 2.27**	95.13 ± 3.36	**98.18 ± 2.43**	**96.61 ± 2.19**	28.491
	CNN4	◇	93.84 ± 3.11	91.41 ± 3.51	96.91 ± 0.96	94.07 ± 2.96	**2.528**
MSS	DT1+DT3	✪	**95.04 ± 1.06**	**93.19 ± 1.87**	**97.27 ± 0.13**	**95.18 ± 0.95**	3.922
	DT1+DT4	⬡	94.07 ± 1.10	92.77 ± 1.69	95.63 ± 0.58	94.17 ± 1.07	**3.840**
	DT2+DT3	⊙	94.98 ± 1.08	93.12 ± 1.88	97.24 ± 0.18	95.12 ± 0.95	4.393
	DT2+DT4	⬠	94.07 ± 1.07	92.77 ± 1.73	95.64 ± 0.68	94.17 ± 1.04	4.382
	CNN1+MN	⊗	96.22 ± 2.29	95.57 ± 3.17	97.13 ± 1.54	96.33 ± 2.21	6.639
	CNN1+IV3	⊠	96.04 ± 2.86	**97.55 ± 1.37**	94.73 ± 5.11	96.04 ± 2.84	14.031
	CNN1+RN50	⊗	**96.78 ± 1.70**	95.15 ± 2.89	**98.79 ± 0.44**	**96.92 ± 1.60**	27.788
	CNN1+CNN4	◈	96.01 ± 1.60	95.40 ± 1.20	96.89 ± 1.93	96.14 ± 1.40	2.872
	CNN2+MN [42]	⊗	96.36 ± 1.85	96.44 ± 2.24	96.41 ± 1.66	96.42 ± 1.83	6.225
	CNN2+IV3	⊠	96.65 ± 1.14	97.34 ± 1.64	96.03 ± 2.33	96.65 ± 1.13	13.749
	CNN2+RN50	⊗	96.75 ± 1.42	96.49 ± 2.37	97.21 ± 1.30	96.83 ± 1.31	27.772
	CNN2+CNN4	◈	95.69 ± 2.06	96.06 ± 1.69	95.39 ± 2.94	95.71 ± 2.80	3.404
	CNN3+MN	⊕	96.40 ± 1.34	95.96 ± 2.39	97.08 ± 0.74	96.50 ± 1.19	6.244
	CNN3+IV3	⊞	96.19 ± 2.82	95.98 ± 3.74	96.59 ± 2.05	96.27 ± 2.80	13.631
	CNN3+RN50	⊕	96.58 ± 1.33	96.21 ± 2.35	97.19 ± 1.87	96.67 ± 1.21	28.116
	CNN3+CNN4	⊕	95.19 ± 2.86	94.43 ± 4.19	96.34 ± 1.31	95.35 ± 2.72	2.717
	MLP+MN	⊗	95.71 ± 2.27	94.70 ± 2.98	97.10 ± 1.67	95.87 ± 2.11	6.479
	MLP+IV3	⊠	96.48 ± 1.91	96.38 ± 1.92	96.78 ± 2.34	96.57 ± 1.83	13.494
	MLP+RN50	⊗	95.64 ± 1.34	94.60 ± 2.87	96.98 ± 1.75	95.73 ± 1.27	28.112
	MLP+CNN4	◈	94.62 ± 2.81	92.20 ± 4.72	98.05 ± 1.75	94.95 ± 2.38	**2.655**
CS	DT1+DT3	✪	**94.42 ± 1.40**	92.64 ± 2.40	96.65 ± 0.21	**94.58 ± 1.21**	2.791
	DT1+DT4	⬡	93.78 ± 1.55	90.93 ± 2.85	97.38 ± 0.32	94.02 ± 1.44	**1.757**
	DT2+DT3	⊙	94.40 ± 1.39	**92.66 ± 2.41**	96.58 ± 0.23	94.56 ± 1.20	3.314
	DT2+DT4	⬠	93.80 ± 1.40	90.77 ± 2.62	**97.60 ± 0.32**	94.04 ± 1.32	2.141
	CNN1+MN	⊗	95.37 ± 2.12	93.92 ± 3.60	97.27 ± 0.95	95.53 ± 1.97	4.050
	CNN1+IV3	⊠	95.84 ± 1.97	95.97 ± 0.57	95.77 ± 4.61	95.80 ± 2.20	7.457
	CNN1+RN50	⊗	**96.68 ± 2.16**	95.50 ± 3.22	98.13 ± 1.13	**96.78 ± 2.09**	19.026
	CNN1+CNN4	◈	94.46 ± 2.52	92.03 ± 3.10	97.47 ± 1.77	94.66 ± 2.40	1.835
	CNN2+MN [42]	⊗	95.28 ± 1.97	94.18 ± 3.45	96.75 ± 1.53	95.41 ± 1.84	3.976
	CNN2+IV3	⊠	95.70 ± 2.39	96.50 ± 0.91	94.96 ± 5.75	95.61 ± 2.73	9.394
	CNN2+RN50	⊗	96.55 ± 2.30	95.21 ± 3.40	98.20 ± 1.17	96.66 ± 2.22	19.278
	CNN2+CNN4	◈	94.31 ± 2.77	92.52 ± 3.08	96.53 ± 2.70	94.47 ± 2.69	2.045
	CNN3+MN	⊕	95.21 ± 2.16	93.40 ± 3.64	97.55 ± 0.86	95.39 ± 1.98	3.960
	CNN3+IV3	⊞	95.90 ± 1.73	95.75 ± 0.76	96.16 ± 4.33	95.89 ± 1.92	7.751
	CNN3+RN50	⊕	96.55 ± 2.22	95.15 ± 3.29	**98.27 ± 1.11**	96.67 ± 2.14	20.022
	CNN3+CNN4	⊕	94.23 ± 2.61	91.40 ± 3.09	97.78 ± 1.77	94.47 ± 2.45	**1.761**
	MLP+MN	⊗	95.14 ± 2.14	93.72 ± 3.66	97.05 ± 0.96	95.31 ± 1.96	4.470
	MLP+IV3	⊠	95.70 ± 2.30	**96.29 ± 0.64**	95.16 ± 5.24	95.63 ± 2.58	9.549
	MLP+RN50	⊗	96.59 ± 2.10	95.29 ± 3.05	98.17 ± 1.17	96.69 ± 2.03	19.864
	MLP+CNN4	◈	94.05 ± 2.87	91.54 ± 3.24	97.19 ± 2.18	94.28 ± 2.73	2.013

5.1 Single-sensor systems photodiode

Among the SSS,PD-configurations, CNN1 has the highest accuracy of 89.84 %, whereas MLP has the lowest of 85.98 %. As explained, the input of SSS,PD only consists of 13 samples. For that few data, CNN1, consisting of a NN with 229 281 parameters, is big. Since only 13 samples are processed, it could be expected that small NN like MLP with 136 parameters would perform accordingly. However, it seems like, complex architectures are beneficial for low-dimensional input data. Out of the NN-SSS,PD-configurations, CNN1 leads to the highest inference time of 0.700 ms and MLP to the lowest of 0.456 ms.

Among the DT-based configurations, DT2, which included automated feature extraction, performs 0.48 % better than DT1, which manually extracted seven features. Because of the slight difference, both systems are competitive. However, since in DT2 more features were extracted, the inference time increases by 0.542 ms.

Comparing the NN- and DT-based SSS,PD-configurations, the best DT-configuration DT2 performs by only 0.09 % worse than the best NN-configuration CNN1. Moreover, DT1 and DT2 perform better than CNN2, CNN3 and MLP with respect to the accuracy. Therefore, we conclude that DT-based methods are competitive with deep-learning-based methods. Deep NN are able to classify without process knowledge. However, classification results of NN are hardly interpretable. In contrast, DT-based methods invest in feature engineering where process knowledge is useful. Therefore, classification results from DT are easier to interpret and additional process information could be found.

5.2 Single-sensor systems high-speed camera

Comparing SSS,HSC with SSS,PD, all SSS,HSC lead to higher accuracies. The reason is that SSS,HSC analyze images that have more information than 13 samples used in SSS,PD. By looking at the SSS,PD symbols (yellow) and the SSS,HSC symbols (red) in Figure 6, the trade-off between accuracy and inference time is visible: On one hand, all SSS,PD have lower inference times; on the other hand, all SSS,HSC have higher accuracies.

Among the NN-based SSS,HSC, RN50 achieves the highest accuracy of 96.50 %. In general, the accuracy of the NN-based configurations increases when the model size increases as well: From CNN4 to MN to IV3 to RN50. However, since RN50 is the largest model with 32 075 393 parameters, the inference time also is the highest with 28.491 ms. The DT-based SSS,HSC perform at the lower end regarding

the accuracy. However, especially DT3 with an accuracy of 93.76 % is competitive with CNN4.

5.3 Multi-sensor systems

MSS mostly result in higher accuracies than the corresponding SSS,HSC. This might lead to the conclusion, that some information in the PD signals is not in the HSC images. On closer inspection, however, the higher the accuracy of the SSS,HSC is, the lower the benefit when using the PD signals additionally. As the benefit of the MSS for complex algorithms is small, we conclude that the main information of the PD signals is included in the HSC images and can be found by using complex processing algorithms. However, when having limited inference times, our results show that the use of both sensors is reasonable.

5.4 Cascaded systems

Every CS is able to outperform its corresponding SSS,HSC with respect to accuracy and inference time (see the red and black symbols and gray lines in Figure 6). By combining two SSS in the proposed cascaded way, it could be that every chunk SSS,PD classifies correctly, SSS,HSC also classifies correctly. The difference would be that SSS,HSC also correctly classifies chunks that SSS,PD cannot. However, this is not the case: Since the accuracies of the CS increase compared with SSS,HSC, SSS,PD sometimes performs better than SSS,HSC. The proposed CS can exploit this effect. Therefore, we conclude that the CS can reduce the inference time with the same or even increased accuracy compared to the corresponding SSS,HSC.

With the help of Figure 6, depending on the conditions of the production process, a suitable system can be chosen. If the inference time cannot be longer than 1 ms, the optimal solution out of all systems would be the CS that combines CNN3 and CNN4 with an accuracy of 92.92 %. This optimal solution can only be seen in Figure 6: It is where the vertical line at an inference time of 1 ms crosses the highest gray line of a CS. Consequently, the evaluations in the second stage so the evaluations of the HSC images of the CS are reduced. Therefore, by a suitable choice of r, our proposed CS leads to the best performance for a given inference time. Analogously, if the inference time is limited to 2 ms, suitable solutions would be the CS that combines CNN1 and MN with an accuracy of 94.40 % or the CS that combines CNN1 and CNN4 with an accuracy of 94.44 %.

The sampling rates of the sensors are shorter than the inference times. This is because the algorithms and the hardware are not optimized yet. Additionally, the input data could be preprocessed more efficiently. For example, several

HSC images could be aggregated into one. Classification then could be performed based on the aggregated image. On one hand, this has the advantage of shorter evaluation times and, on the other hand, a larger neighborhood is considered. The latter is useful since weld defects are sometimes visible in several consecutive images.

In general, CS are worthwhile, especially for short inference times. For longer inference times MSS outperform. However, from a hardware perspective, CS could be performed on two separate configurations at each sensor where only one bit has to be exchanged between the two configurations, indicating if the SSS,PD is sure or not. In contrast, when performing MSS, all data have to be transferred to a common hardware to process them together.

5.5 Friedman and Nemenyi test

To compare the performance of the different SSS,PD and SSS,HSC models, a Friedman test with a significance level of 0.05 was performed. The Friedman test shows that there is a significant difference between the SSS,PD and SSS,HSC models, respectively. Therefore, in a next step, a Nemenyi test with the same significance level as the Friedman test was performed. This led to the result that for the SSS,PD models there is a significant difference between configuration MLP and CNN1. For the SSS,HSC models, there is a significant difference between DT4 and IV3, DT4 and RN50, and DT3 and RN50.

6 Summary and outlook

A confidence-based cascaded system (CS) as a quality monitoring approach for a laser welding process is presented. The key idea of the CS is that not all data are analyzed to obtain the quality weld, but only selected ones. The CS is compared with state-of-the-art quality monitoring approaches, namely single-sensor systems (SSS) and multi-sensor systems (MSS). For every approach, various algorithms consisting of feature engineering and decision trees (DT) or neural networks (NN) only are compared in terms of the accuracy and the inference time since fast but still precise quality monitoring is needed in today's quality monitoring of welds. Thereby, all CS are able to outperform SSS in terms of accuracy and inference time. Depending on the conditions of a production process an optimal CS can be chosen. Thereby, especially for short inference times the CS is the best choice among the approaches.

Regarding the algorithms, classical machine learning using feature engineering and DT results in competitive results to NN. Thereby, results coming from classical machine learning are easier to interpret but can lead to time investment in feature engineering. Moreover, compared to MSS, the presented CS has the advantage, that data of different sensors do not have to be transferred to common hardware. However, when transferred to common hardware, the MSS could be executed in parallel, while the CS must be executed sequentially.

Further work can deal with a general cascaded system [47] for quality monitoring, which arbitrarily combines any sensors, to obtain a quality assessment. As multi-modal systems work well under information fusion, additional non-optical sensors like acoustical sensors could be added. Moreover, further optical sensors like microscopy could be used after the welding process. While PD and HSC provide information during the welding process like about the keyhole and the molten metal, microscopy images of the solidified weld seam could be used additionally. Furthermore, to improve the CS, classifiers in lower stages could only be trained on the subset of data that are not already confidently classified by the higher stages. However, this would require to train a new classifier for each r value. Additionally, sensors could deliver conflicting data. Therefore, it would be interesting to analyze what proportion of the error rate of a CS is due to an incorrect and overconfident decision of the first stage, meaning the second stage would have corrected that misclassification, and what proportion is also misclassified by the second stage. Moreover, the quality monitoring approaches can be extended to distinguish not only between anomaly and reference but also between different anomalies.

References

[1] G. Chen, L. Mei, M. Zhang, Y. Zhang, and Z. Wang, "Research on key influence factors of laser overlap welding of automobile body galvanized steel," *Opt Laser. Technol.*, vol. 45, pp. 726−733, 2013.

[2] K. M. Hong and Y. C. Shin, "Prospects of laser welding technology in the automotive industry: a review," *J. Mater. Process. Technol.*, vol. 245, pp. 46−69, 2017.

[3] K. Abderrazak, W. Ben Salem, H. Mhiri, P. Bournot, and M. Autric, "Nd: YAG laser welding of AZ91 magnesium alloy for aerospace industries," *Metall. Mater. Trans. B*, vol. 40, pp. 54−61, 2009.

[4] K. Haug and G. Pritschow, "Robust laser-stripe sensor for automated weld-seam-tracking in the shipbuilding industry," in *IECON'98. Proceedings of the 24th Annual Conference of the IEEE Industrial Electronics Society (Cat. No. 98CH36200)*, vol. 2, 1998, pp. 1236−1241.

[5] M. M. Atabaki, N. Yazdian, J. Ma, and R. Kovacevic, "High power laser welding of think steel plates in a horizontal butt joint configuration," *Opt Laser. Technol.*, vol. 83, pp. 1−12, 2016.

[6] S. Pang, X. Chen, J. Zhou, X. Shao, and C. Wang, "3D transient multiphase model for keyhole, vapor plume, and weld pool dynamics in laser welding including the ambient pressure effect," *Opt. Lasers Eng.*, vol. 74, pp. 47−58, 2015.

[7] Y. Zhang, F. Li, Z. Liang, Y. Ying, Q. Lin, and H. Wei, "Correlation analysis of penetration based on keyhole and plasma plume in laser welding," *J. Mater. Process. Technol.*, vol. 256, pp. 1−12, 2018.

[8] C. Alippi, P. Braione, V. Piuri, and F. Scotti, "A methodological approach to multisensor classification for innovative laser material processing units," in *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference (I2MTC)*, vol. 3, 2001, pp. 1762−1767.

[9] K. Hellera, S. Kesslera, F. Dorscha, P. Berger, and T. Graf, "Robust "false friend" detection via thermographic imaging," in *Lasers in Manufacturing Conference 2015*, 2015.

[10] J. Powell, T. Ilar, J. Frostevarg, et al., "Weld root instabilities in fiber laser welding," *J. Laser Appl.*, vol. 27, no. S2, p. S29008, 2015.

[11] J. Frostevarg and A. F. H. Kaplan, "Undercuts in laser arc hybrid welding," *Phys. Procedia*, vol. 56, pp. 663−672, 2014.

[12] A. G. Paleocrassas and J. F. Tu, "Inherent instability investigation for low speed laser welding of aluminum using a single-mode fiber laser," *J. Mater. Process. Technol.*, vol. 210, no. 10, pp. 1411−1418, 2010.

[13] A. Molino, M. Martina, F. Vacca, et al., "FPGA implementation of time−frequency analysis algorithms for laser welding monitoring," *Microprocess. Microsyst.*, vol. 33, no. 3, pp. 179−190, 2009.

[14] S. S. Rodil, R. A. Gómez, J. M. Bernández, F. Rodríguez, L. J. Miguel, and J. R. Perán, "Laser welding defects detection in automotive industry based on radiation and spectroscopical measurements," *Int. J. Adv. Des. Manuf. Technol.*, vol. 49, pp. 133−145, 2010.

[15] G. Chianese, P. Franciosa, J. Nolte, D. Ceglarek, and S. Patalano, "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. Manuf. Sci. Eng.*, vol. 144, no. 7, p. 071004, 2022.

[16] F. Kong, J. Ma, B. Carlson, and R. Kovacevic, "Real-time monitoring of laser welding of galvanized high strength steel in lap joint configuration," *Opt Laser. Technol.*, vol. 44, pp. 2186−2196, 2012.

[17] P. B. García-Allende, J. Mirapeix, O. M. Conde, A. Cobo, and J. M. López-Higuera, "Spectral processing technique based on feature selection and artificial neural networks for arc-welding quality monitoring," *NDT&E Int.*, vol. 42, no. 1, pp. 56−63, 2009.

[18] M. Thornton, L. Han, and M. Shergold, "Progress in NDT of resistance spot welding of aluminium using ultrasonic C-scan," *NDT&E Int.*, vol. 48, pp. 30−38, 2012.

[19] K. Wasmer, T. Le-Quang, B. Meylan, et al., "Laser processing quality monitoring by combining acoustic emission and machine learning: a high-speed X-ray imaging approach," *Procedia CIRP*, vol. 74, pp. 654−658, 2018.

[20] S. Shevchik, T. Le-Quang, B. Meylan, et al., "Supervised deep learning for real-time quality monitoring of laser welding with X-ray radiographic guidance," *Sci. Rep.*, vol. 10, no. 1, p. 3389, 2020.

[21] M. Baader, A. Mayr, T. Raffin, J. Selzam, A. Kühl, and J. Franke, "Potentials of optical coherence tomography for process monitoring in laser welding of hairpin windings," in *11th International Electric Drives Production Conference (EDPC)*, 2021, pp. 1−10.

[22] S. Tsukamoto, "High speed imaging technique Part 2 − high speed imaging of power beam welding phenomena," *Sci. Technol. Weld. Joining*, vol. 16, no. 1, pp. 44−55, 2011.

[23] M. Jäger, S. Humbert, and F. A. Hamprecht, "Sputter tracking for the automatic monitoring of industrial laser-welding processes," *IEEE Trans. Ind. Electron.*, vol. 55, pp. 2177−2184, 2008.

[24] M. Jäger and F. A. Hamprecht, "Principal component imagery for the quality monitoring of dynamic laser welding processes," *IEEE Trans. Ind. Electron.*, vol. 56, pp. 1307−1313, 2009.

[25] J. Vater, M. Pollach, C. Lenz, D. Winkle, and A. Knoll, "Quality control and fault classification of laser welded hairpins in electrical motors," in *28th European Signal Processing Conference (EUSIPCO)*, 2017, pp. 1377−1381.

[26] B. Zhou, T. Pychynski, M. Reischl, E. Kharlamov, and R. Mikut, "Machine learning with domain knowledge for predictive quality monitoring in resistance spot welding," *J. Intell. Manuf*, vol. 33, no. 4, pp. 1139−1163, 2022.

[27] E. B. Schwarz, F. Bleier, F. Guenter, R. Mikut, and J. P. Bergmann, "Improving process monitoring of ultrasonic metal welding using classical machine learning methods and process-informed time series evaluation," *J. Manuf. Process.*, vol. 77, pp. 54−62, 2022.

[28] T. S. Yun, K. J. Sim, and H. J. Kim, "Support vector machine-based inspection of solder joints using circular illumination," *Electron. Lett.*, vol. 36, no. 11, p. 1, 2000.

[29] D. You, X. Gao, and S. Katayama, "Multisensor fusion system for monitoring high-power disk laser welding using support vector machine," *IEEE Trans. Ind. Inform.*, vol. 10, no. 2, pp. 1285−1295, 2014.

[30] Z. Zhang and S. Chen, "Real-time seam penetration identification in arc welding based on fusion of sound, voltage and spectrum signals," *J. Intell. Manuf*, vol. 28, no. 1, pp. 207−218, 2017.

[31] X. Hongwei, Z. Xianmin, K. Yongcong, and O. Gaofei, "Solder joint inspection method for chip component using improved AdaBoost

and decision tree," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 1, no. 12, pp. 2018−2027, 2011.

[32] H. Wu, "Solder joint defect classification based on ensemble learning," *Solder. Surf. Mt. Technol.*, vol. 29, no. 3, pp. 164−170, 2017.

[33] C. Knaak, U. Thombansen, P. Abels, and M. Kröger, "Machine learning as a comparative tool to determine the relevance of signal features in laser welding," *Procedia CIRP*, vol. 74, pp. 623−627, 2018.

[34] H. Wu, X. Zhang, H. Xie, Y. Kuang, and G. Ouyang, "Classification of solder joint using feature selection based on Bayes and support vector machine," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 3, no. 3, pp. 516−522, 2013.

[35] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84−90, 2017.

[36] Y. Yang, L. Pan, J. Ma, et al., "A high-performance deep learning algorithm for the automated optical inspection of laser welding," *Appl. Sci.*, vol. 10, no. 3, p. 933, 2020.

[37] Y. Yang, R. Yang, L. Pan, et al., "Real-time monitoring of high-power disk laser welding statuses based on deep learning framework," *J. Intell. Manuf*, vol. 31, pp. 799−814, 2020.

[38] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818−2826.

[39] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations (ICLR)*, 2015.

[40] A. Howard, M. Sandler, G. Chu, et al., "Searching for MobileNetV3," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2017, pp. 1314−1324.

[41] P. Stritt, M. Boley, A. Heider, et al., "Comprehensive process monitoring for laser welding process optimization," in *Proc. SPIE 9741, High-Power Laser Materials Processing: Lasers, Beam Delivery, Diagnostics, and Applications V*, vol. 9741, 2016, pp. 193−202.

[42] D. You, X. Gao, and S. Katayama, "Data-driven based analyzing and modeling of MIMO laser welding process by integration of six advanced sensors," *Int. J. Adv. Des. Manuf. Technol.*, vol. 82, nos. 5−8, pp. 1127−1139, 2016.

[43] Y. Zhang, D. You, X. Gao, N. Zhang, and P. P. Gao, "Welding defects detection based on deep learning with multiple optical sensors during disk laser welding of thick plates," *J. Manuf. Syst.*, vol. 51, pp. 87−94, 2019.

[44] T.-H. Kim, T.-H. Cho, Y. S. Moon, and S. H. Park, "Visual inspection system for the classification of solder joints," *Pattern Recognit.*, vol. 32, no. 4, pp. 565−575, 1999.

[45] S.-C. Lin, C. H. Chou, and C.-H. Su, "A development of visual inspection system for surface mounted devices on printed circuit board," in *IECON 2007-33rd Annual Conference of the IEEE Industrial Electronics Society*, 2007, pp. 2440−2445.

[46] K. Y. Chan, K. F. C. Yiu, H.-K. Lam, and B. W. Wong, "Ball bonding inspections using a conjoint framework with machine learning and human judgement," *Appl. Soft Comput.*, vol. 102, p. 107115, 2021.

[47] P. M. Dold, F. Bleier, M. Boley, and R. Mikut, "Multi-stage inspection of laser welding defects using machine learning," in *Proceedings 32. Workshop Computational Intelligence*, vol. 1, 2022, pp. 31−52.

[48] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[49] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, "Time series feature extraction on basis of scalable hypothesis tests (tsfresh − a Python package)," *Neurocomputing*, vol. 307, pp. 72−77, 2018.

[50] A. G. Howard, M. Zhu, B. Chen, et al., "Efficient convolutional neural networks for mobile vision applications," *arXiv:1704.04861*, 2017.

[51] F. Pedregosa, G. Varoquaux, A. Gramfort, et al., "Scikit-learn: machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825−2830, 2011.

[52] F. Chollet, "Keras," in *GitHub*, 2015. Available at: https://github.com/fchollet/keras.

[53] M. Abadi, P. Barham, J. Chen, et al., "TensorFlow: a system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016, pp. 265−283.

[54] A. Brand, L. Allen, M. Altman, M. Hlava, and J. Scott, "Beyond authorship: attribution, contribution, collaboration, and credit," *Learn. Publ.*, vol. 28, no. 2, pp. 151−155, 2015.

Bionotes

Patricia M. Dold
Bosch Research, Robert Bosch GmbH, Robert-Bosch-Campus 1, 71272 Renningen, Germany; and Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany
patricia.dold@de.bosch.com

Patricia M. Dold received the B.S. degree and the M.S. degree in electrical engineering and information technology from the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2019 and 2022, respectively. In 2020, she spent one semester at the Polytechnic University of Valencia (UPV), Valencia, Spain. Currently, her research at Bosch Research, Renningen, Germany, and the Institute for Automation and Applied Informatics (IAI), KIT, Eggenstein-Leopoldshafen, Germany, focuses on distributed multi-sensor quality monitoring of laser-based processes for intelligent production.

Fabian Bleier
Bosch Research, Robert Bosch GmbH, Robert-Bosch-Campus 1, 71272 Renningen, Germany

Dr. Fabian Bleier is a research engineer in the department of Advanced Production Technologies at Bosch Research in Renningen. His research focuses on the integration of data analytics in manufacturing environments from edge to cloud with a special interest in real-time capable data analytics on resource-restricted devices.

Meiko Boley
Bosch Research, Robert Bosch GmbH,
Robert-Bosch-Campus 1, 71272 Renningen,
Germany

Dr. Meiko Boley is a research engineer and works at Bosch Research in Renningen. He wrote his Ph.D. about using optical coherence tomography in laser welding. His current work is centered around process monitoring of laser-based processes in production.

Ralf Mikut
Institute for Automation and Applied
Informatics (IAI), Karlsruhe Institute of
Technology (KIT),
Hermann-von-Helmholtz-Platz 1, 76344
Eggenstein-Leopoldshafen, Germany

apl. Prof. Dr. Ralf Mikut received the Dipl.-Ing. degree in automatic control from the University of Technology, Dresden, Germany, in 1994, and the Ph.D. degree in mechanical engineering from the University of Karlsruhe, Karlsruhe, Germany, in 1999. Since 2011, he has been an Adjunct Professor at the Faculty of Mechanical Engineering and the head of the research field "Automated Image and Data Analysis". He is leading the research group "Machine Learning for Time Series and Images" at the Institute for Automation and Applied Informatics of the Karlsruhe Institute of Technology (KIT), Germany. His current research interests include machine learning, image processing, life science applications and smart grids.