Multi-stage Inspection of Laser Welding Defects using Machine Learning

Patricia M. Dold[1,2], Fabian Bleier[1], Meiko Boley[1], Ralf Mikut[2]

[1] Bosch Research
Robert Bosch GmbH
Robert-Bosch-Campus 1, 71272 Renningen
E-Mail: patricia.dold@de.bosch.com
[2]Institute for Automation and Applied Informatics
Karlsruhe Institute of Technology
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen

Abstract

As welding processes become faster and components consist of many more welds compared to previous applications, there is a need for fast but still precise quality inspection. The aim of this paper is to compare already existing approaches, namely single-sensor systems (SSS) and multi-sensor systems (MSS) with a proposed cascaded system (CS). The introduced CS is characterized by the fact that not all available data are analyzed, but only cleverly selected ones. The different approaches consisting of neural networks are compared in terms of their accuracy and computational effort. The data are recorded from scratch and include two common sensor systems for quality control, namely a photodiode (PD) and a high-speed camera (HSC). As a result, when the CS makes half of the final decisions based on a SSS with PD signals and the other half based on a SSS with HSC images, the estimated computational effort is reduced by almost 50% compared to the SSS with HSC images as input. At the same time, the accuracy decreases only by 0.25% to 95.96%. Additionally, based on the CS, a general cascaded system (GCS) for quality inspection is proposed.

DOI: 10.58895/ksp/1000151141-3 erschienen in:
Proceedings – 32. Workshop Computational Intelligence: Berlin, 1. - 2. Dezember 2022
DOI: 10.58895/ksp/1000151141 | https://www.ksp.kit.edu/site/books/m/10.58895/ksp/1000151141/

1 Introduction

Laser welding is widely used in automotive, aerospace or shipbuilding industries and is considered a key technology in manufacturing [1, 2, 3]. Advantageous compared to other joining techniques is its ability for precise and fast welding. Unfortunately, laser welding, in which a workpiece is melted, vaporized and solidified, often is a challenging process [4] and, hence, welding defects occur. Especially in the electric drive train, some components consist of several hundred laser-welded elements, and a single welding defect can lead to the failure of the entire component. In order to detect defective elements in time, there is a desire in the industry for quality monitoring.

The most important sensor methods for examining weld quality include photodiodes (PD) [5, 6, 7] or spectrometers [8, 9] due to their simple structure and low cost. Other methods to provide information about spatter, keyhole, weld pool or plasma are ultraviolet sensors [10], X-rays [11, 12], microscopy, optical coherent tomography (OCT) [13] or high-speed cameras (HSC) [14, 15, 16, 17]. Accordingly, compared to typical quality monitoring in resistance welding [18] or ultrasonic welding [19], there are more potential input variables in laser welding.

The signals acquired by the different sensors are analyzed using signal or image processing algorithms. Thereby, data-driven process monitoring has been implemented by applying machine learning methods such as support vector machines (SVM) [20], decision trees [21], random forest algorithms [22] or Bayes classifiers [23]. Recently, deep learning has achieved great success in image recognition and classification [24] and thus has been applied to weld defect inspection, especially convolutional neural networks (CNN).

Some researchers use a single sensor type to study a specific mechanism of the welding process: In [6] the optical intensity captured by a PD when welding defects occur is analyzed. [17] does quality assessment of welds based on HSC or OCT data using deep neural networks like Inception-v3 [25]. Different CNNs like AlexNet, VGG-16 or MobileNetV3-Large [26] are used in [27] for automated optical inspection of laser welding.

However, information captured by one sensor is not sufficient for holistic quality assessment [28]. A combination of different sensor types, on the other hand, provides a more comprehensive description of the welding process [11]. In [29] different sensor types are used, including an ultraviolet-visible band visual sensor system, a spectrometer or a PD in order to detect three different weld defects during high power disk laser welding using neural networks. In [30] also several sensors in order to predict the welding quality are applied using machine learning methods.

A quality-monitoring system whose processing time is not longer than its cyclic time would be optimal. Yet, complex multi-sensor systems and processing algorithms can result in quality-monitoring systems that are not used in production due to long evaluation times. For this reason, the present paper proposes a cascaded system (CS) with the aim of fast but still precise quality inspection. The system follows a multi-stage structure: The first level of inspection has time series as input, with the advantage of a high clock rate as well as low memory requirements. This stage already safely classifies some welds; in areas of uncertainty, on the other hand, the next more complex stage with image data as input takes over in order to make a final decision. Quality control based on two-stages has already been applied to other use cases [23, 31, 32, 33]. In [23] a two-stage classifier for solder joint inspection has been proposed. After feature selection based on the algorithm of Bayes, each solder joint is classified by its qualification. If the solder joint fails in a qualification test, it is classified based on a SVM. Moreover, in [33] an inspection system for ball bonding is incorporated using CNNs or SVMs, where human judgment is only used when the detection uncertainty is below a threshold. The present paper deals with a cascaded system for quality control. Thereby, it links to already existing ideas of two-stage process monitoring in the literature. To show the added value of the system presented in this paper, it is compared to a single-sensor and multi-sensor systems. It lays the foundation for further cascaded decision-making systems in quality control.

2 Data Set

Since public data sets are usually not available for industrial research, especially for welding applications, the data used in this paper were acquired from scratch in the laboratory. Photodiode signals and associated high-speed camera images captured during the welding of 60 pairs of metal plates are considered. Subsequently, the recorded data were preprocessed so that they can be used by the various quality monitoring approaches considered in this paper.

2.1 Experimental Setup

In the considered laser welding process, two metal plates with a thickness of 75 µm were welded together with a power of 250 W in a commercial setup. The plates were placed on top of each other, clamped and welded in a rectangular geometry. Figure 1a illustrates the welding process of the two metal plates in cross-section. Figure 1b shows the rectangular welding path. The metal plate pair is viewed from above, which means that the plates are on top of each other in the image plane. Different anomalies such as spatters or gaps were manually inserted into the welding process. Of 60 welded plates, anomalies were inserted in 51, whereas 9 were welded under reference conditions. The dashed line indicates the position of anomalies. The initial position of the laser beam was $(x,y) = (0,0)$. It was then deflected to position $(x_{\text{on}}, y_{\text{on}})$ where the laser starts to weld the geometry and ends in position $(x_{\text{off}}, y_{\text{off}})$. Finally, the laser beam returned to its initial position and was ready for the next weld. In total, the welding process for a pair of metal plates took 338 ms.

Figure 2 shows the experimental setup. The laser optics were located above the metal plates at $(x,y,z) = (0,0,40\,\text{cm})$ and directed the laser beam to the rectangular path using two mirrors. A photodiode (PD) measured light with a wavelength of about 300–950 nm obtained during the welding process in the area of the weld pool. Accordingly, at each sampling time, the voltage of the PD amplifier was recorded. Furthermore, HSC images were taken during the process. The PD had a sampling rate of 250 kHz, whereas the HSC had one of 20 kHz. The advantage of the PD is the higher sampling rate as this allows

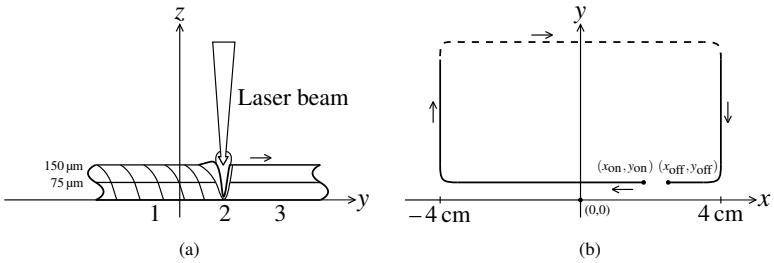

Figure 1: Laser welding process. (a) shows the two metal plates in cross-section during the welding process. There are three areas: 1 solidified melt, 2 weld pool and 3 unwelded metal plates. In (b) the rectangular welding path is indicated. Dashed lines indicate where anomalies were introduced in the laboratory experiment. Arrows indicate the welding direction.

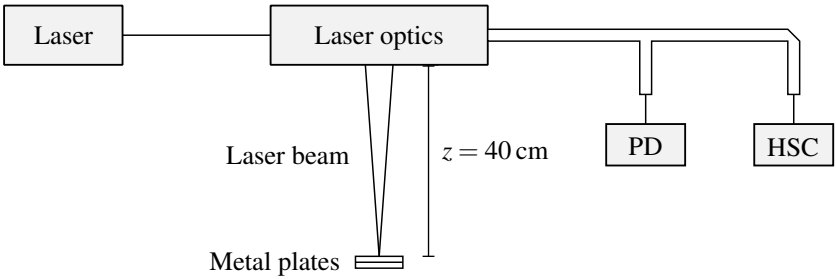

Figure 2: Experimental setup of the laser welding process of two metal plates. The two used sensors for data acquisition are a photodiode (PD) and a high-speed camera (HSC).

the detection of shorter anomalies compared to the HSC. Moreover, faster data processing is possible due to the smaller amount of raw data compared to the HSC. However, the HSC images provide information which is not available in the PD signals, e.g. information about the current geometry of the interaction zone. Figure 3 illustrates the sampling rates and the data of both sensors. The gray lines on the x-axis correspond to the sampling times of the PD and have a distance of 4 µs each. The black larger lines on the axis indicate the sampling times of the HSC, which are 50 µs apart. Every 12.5 samples of the PD are followed by a HSC gray scale image. In addition to the sampling rates, the PD and HSC data are visualized. In the range from 500–700 µs data of reference welds are shown. At $t = 250\,000$ µs an anomaly occurs, more precisely, the process spatters. On closer inspection, a slightly different interaction zone

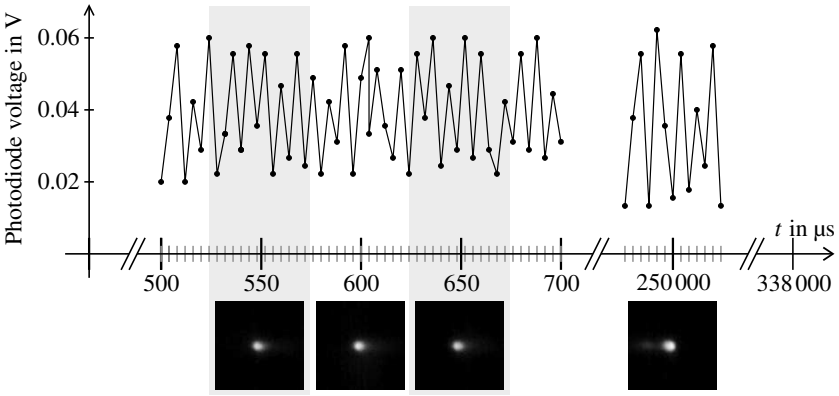

Figure 3: Illustration of the PD and HSC data and sampling rates.

geometry and a splash on the left side can be seen compared to the reference HSC images.

2.2 Preprocessing

For the quality-monitoring approaches considered in this paper, 13 samples of the PD were assigned to each image of the HSC. The assignment of the PD samples to the corresponding HSC image is visualized in Figure 3 by the gray background. Besides, the gray scale HSC images were cropped to a section of 100×100 pixels and scaled to a value range of [0,1].

A label was manually assigned to each image with the corresponding 13 PD samples, distinguishing between reference and anomaly. By section-wise labeling of the weld seams, a localization of reference and anomaly on the welding path is possible. This would not be possible when labeling a pair of plates as a whole. An anomaly refers to those locations, where an abnormality like a gap or spatters have been introduced. Reference is defined as those places where neither an abnormality is provoked nor is visible in the recorded PD signals or in the HSC images. There are places, where no intentional defects were introduced, but abnormalities are visible in the signal. One cause is sporadically occurring errors. Moreover, micrograph analyses show that

Table 1: Number of samples of the recorded data set. During training, data from 48 metal plates (80%) is used; during testing, data from 12 plates (20%) is used. The number of samples n_{DA} generated by data augmentation (DA) is also given. The values given indicate the average sample numbers over the 5 folds of the used cross-validation.

Label	n_{train}	n_{test}	n_{total}	n_{DA}
Reference	148 425	37 106	185 531	1 576 738
Anomaly	152 777	38 194	190 971	1 623 262
Total	301 202	75 300	376 502	3 200 000

anomalies may be present in the weld at that places. These positions are left out in the following because the correct label cannot be determined and would therefore confuse the process monitoring systems. Of the 60 metal plate units, the proportion of such identified locations is 7.19%.

Table 1 shows the number of samples of the data set used in this paper. During training, data from 48 metal plates were used and during testing, data from 12. For the evaluation of the models, a 5-fold cross-validation was applied. Thereby, in every fold the data of 12 metal plates were left out. Since different numbers of samples were left out on each metal plate, the number of samples varied in every fold. The average number of training samples over the 5 folds is given by n_{train}; the average number of test samples by n_{test} and both together as n_{total}. In order to teach the models invariances and robustness properties, a data augmentation on the PD signals as well as on the HSC images was used. The PD signals were reflected horizontally. The HSC images were rotated and reflected. This data augmentation is essential for the HSC images since, as shown in Figure 1b, anomalies were only inserted on one side of the welded geometry. If not applying this data augmentation, the classification algorithms could learn, that anomalies only take place in one direction, which is not the case in real processes.

3 Different Approaches for Quality Monitoring

Three different approaches, namely single-sensor systems (SSS), multi-sensor systems (MSS) and cascaded systems (CS) are considered. Before the approaches are introduced, the formal definitions of the analyzed data are defined:

For each sample, the time series of the PD signal corresponding to a HSC image is defined as

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

where $i \in \{1, \dots, 376\,502\}$ indicates the considered sample number. The images of the HSC are defined as

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

With the label $Y_i \in \{0, 1\}$, where 0 stands for anomaly and 1 for reference, the data set S consists of 376 502 triples according to

$$S = \{(X_{\mathrm{PD},1}, X_{\mathrm{HSC},1}, Y_1), \dots, (X_{\mathrm{PD},376\,502}, X_{\mathrm{HSC},376\,502}, Y_{376\,502})\}. \tag{3}$$

3.1 Single-Sensor and Multi-Sensor System

A single-sensor system (SSS) performs process monitoring based on data coming from one sensor. This can be represented by the function

$$f_{\mathrm{SSS}} \colon \mathbb{R}^{a \times b} \to \{0,1\} \colon X_i \mapsto Y_i, \quad a, b \in \mathbb{R}. \tag{4}$$

Figure 4 illustrates SSSs based on PD signals or HSC images, where $\hat{f}_{SSS}$ indicates an optimized model. It could be any classical machine learning algorithm

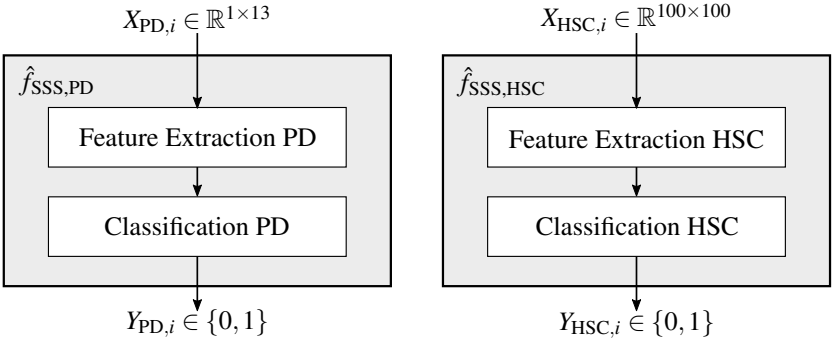

Figure 4: Two single-sensor systems (SSS). Thereby, $\hat{f}_{\mathrm{SSS,PD}}$ and $\hat{f}_{\mathrm{SSS,HSC}}$ are the prediction models based on PD samples $X_{\mathrm{PD},i}$ or HSC samples $X_{\mathrm{HSC},i}$ as input. $Y_{\mathrm{PD},i}$ and $Y_{\mathrm{HSC},i}$ are the predicted labels, which deviate from Y_i in case of misclassification.

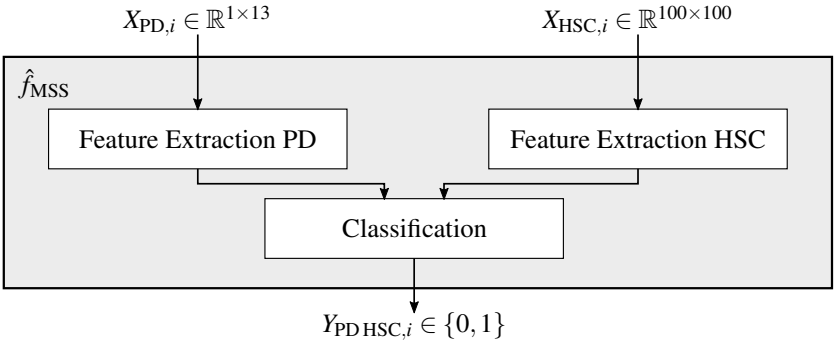

Figure 5: Multi-sensor system (MSS). The prediction model $\hat{f}_{\mathrm{MSS}}$ has PD samples $X_{\mathrm{PD},i}$ and HSC samples $X_{\mathrm{HSC},i}$ as input and predicts $Y_{\mathrm{PD\,HSC},i}$.

or a neural network with convolutional layers for feature extraction and fully-connected layers for decision-making. In the case of a neural network, the output is first a real number, which is then binarized for the final decision. Those predicted labels differ from Y_i in the case of misclassification.

A multi-sensor system (MSS) uses data from multiple sensors for process monitoring. A MSS consisting of PD signals and HSC images as inputs can be expressed as

$$f_{\mathrm{MSS}} : \mathbb{R}^{1\times 13} \times \mathbb{R}^{100\times 100} \to \{0,1\} : X_{\mathrm{PD},i} \times X_{\mathrm{HSC},i} \mapsto Y_i. \tag{5}$$

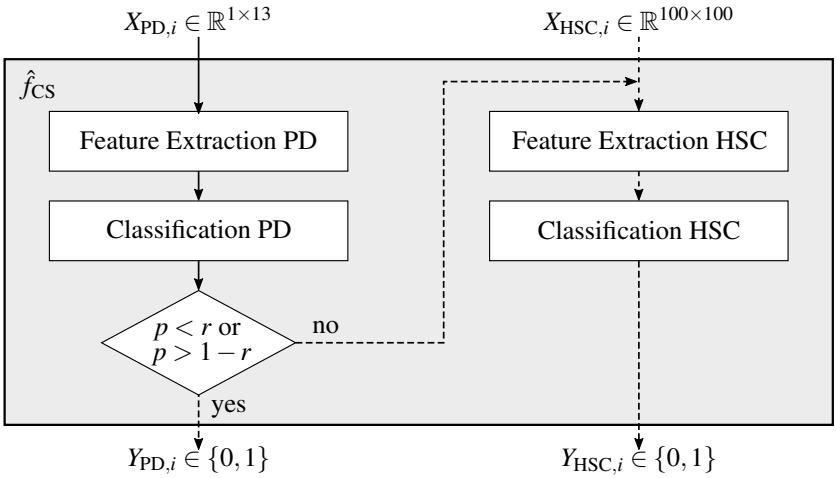

Figure 6: Cascaded system (CS). Depending on p, the prediction model $\hat{f}_{\mathrm{CS}}$ decides either for $Y_{\mathrm{PD},i}$ or $Y_{\mathrm{HSC},i}$.

Figure 5 represents a MSS based on PD signals and HSC images. The advantage of a MSS compared to a SSS is a more holistic quality control. However, the computational effort of the MSS is greater than of the SSS, which is disadvantageous for industrial quality monitoring of fast laser processes.

3.2 Cascaded System

A cascaded system (CS) offers the possibility to use several sensors as a MSS does. In contrast to MSS, on the other hand, it is characterized by the fact that not all available data are analyzed to get the quality-weld condition; only cleverly selected data are evaluated. Figure 6 shows the two-stage CS used in this paper. Let $p \in [0,1]$ be the output of a classifier which makes quality assessment 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 to be. Let $r \in (0, 0.5)$ be a fixed threshold. A decision of the classifier is evaluated as certain if $p < r$ or $p > 1-r$. If the first condition is satisfied, the classifier is sure that it is an anomaly; if the second is satisfied, it is certain that it is a reference. If the classifier is sure, the result

is accepted; if not, a final quality decision is made in a next step based on the HSC data. Formally, this can be expressed by

$$f_{\mathrm{CS}}: \begin{cases} f_{\mathrm{SSS,PD}}, & \text{if} \quad p < r \text{ or } p > 1-r \\ f_{\mathrm{SSS,HSC}}, & \text{otherwise.} \end{cases} \tag{6}$$

4 Network Structures and Training Details

The four considered prediction models $\hat{f}_{\mathrm{SSS,PD}}$, $\hat{f}_{\mathrm{SSS,HSC}}$, $\hat{f}_{\mathrm{MSS}}$ and $\hat{f}_{\mathrm{CS}}$ were built by combining four blocks, namely Feature Extraction PD, Feature Extraction HSC, Classification PD and Classification HSC. The blocks consist of neural network architectures, which are described in detail later. Building the four models based on the same blocks provides the advantage of better performance comparison. $\hat{f}_{\mathrm{SSS,PD}}$ and $\hat{f}_{\mathrm{SSS,HSC}}$ use the combination of two blocks each according to Figure 4. $\hat{f}_{\mathrm{MSS}}$ was constructed according to Figure 5 using Classification HSC as classification block. $\hat{f}_{\mathrm{CS}}$ uses the combination of the four neural network blocks according to Figure 6.

The Feature Extraction PD block is based on [34], where different deep neural networks for time-series classification are proposed. It consists of four 1D convolution layers with 8, 16, 16 and 8 filters respectively and a kernel size of 3. Each convolution layer is followed by batch normalization and a ReLU as activation function. The Feature Extraction HSC block is based on MobileNet [35]. It consists of the MobileNet architecture until the last depthwise separable convolution layer. Pretrained weights based on ImageNet were used. The Classification PD block has three fully-connected layers with 16, 8 and 4 neurons followed by the single output gained with a sigmoid activation function for classification. Every fully-connected layer has a ReLU as activation function followed by a dropout layer with a rate of 0.5. The Classification HSC block consists of the same structure as Classification PD but with the fully-connected layers having 256, 256 and 128 neurons each.

During training of the models, 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 $\beta_2 = 0.99$ were used. The models were trained with a batch

size of 32 with 100 steps per epoch for 1000 epochs. All models were trained with the same seed, meaning that every model is trained with the exact same augmented images or time series respectively in the exact same order for better performance comparison. Additionally, 5-fold cross-validation was used to get the final results. The models were implemented in Python using Keras and Tensorflow. The training processes ran on a NVIDIA A40 GPU card.

5 Results and Discussion

In the following, the performances of the four models $\hat{f}_{\text{SSS,PD}}$, $\hat{f}_{\text{SSS,HSC}}$, $\hat{f}_{\text{MSS}}$ and $\hat{f}_{\text{CS}}$ are compared with respect to the accuracy and the number of parameters as a rough indicator for the computational effort. Let $n_{\overline{m}}$ be the number of samples where the output of $\hat{f}_{\text{SSS,PD}}$ before binarization is greater than $1-r$ or smaller than r, thus being the number of samples, where $\hat{f}_{\text{SSS,PD}}$ is certain. Let n_m be the number of all remaining samples, thus the number of samples, where $\hat{f}_{\text{SSS,HSC}}$ decides. $A_{\text{SSS,PD}\overline{m}}$ and $A_{\text{SSS,HSC}m}$ give the proportion of correctly classified samples, respectively. The accuracy A_{CS} of the cascaded system (CS) over all samples depending on the previously introduced threshold $r \in (0, 0.5)$ is thereby calculated by

$$A_{\text{CS}}(r) = \frac{n_{\overline{m}}(r) \cdot A_{\text{SSS,PD}\overline{m}}(r) + n_m(r) \cdot A_{\text{SSS,HSC}m}(r)}{n_{\overline{m}}(r) + n_m(r)}. \tag{7}$$

The accuracies $A_{\text{SSS,PD}}$, $A_{\text{SSS,HSC}}$, A_{MSS} and A_{CS} of the four models gained through 5-fold cross-validation are shown as black lines in Figure 7. Comparison of the accuracies of the four systems indicates that the SSS with the PD signals as input has the lowest. This is not surprising since in comparison to the SSS with the HSC images as input, only 13 samples of a time series are available at any time. The MSS having the highest accuracy can be explained by the fact that data of two sensors are analyzed simultaneously, and therefore a more holistic quality assessment is possible than with a single sensor. Thus, it can be inferred that there is information in the PD signal that is not captured by the HSC images.

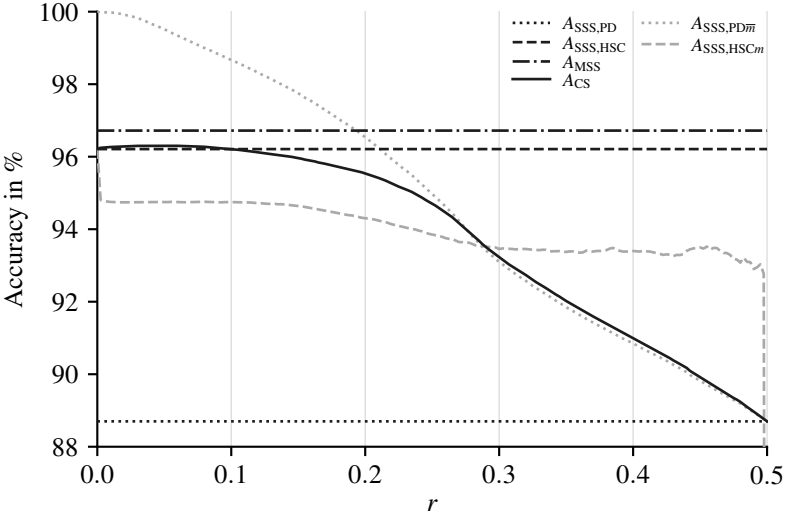

Figure 7: Accuracies. Black lines show the accuracies of the four models. The concrete accuracies of the models which are independent of r are $A_{\mathrm{SSS,PD}} = 88.70\%$, $A_{\mathrm{SSS,HSC}} = 96.21\%$ and $A_{\mathrm{MSS}} = 96.72\%$. At $r = 0.053$, A_{CS} has its maximum of 96.31%. Gray lines indicate the accuracies $A_{\mathrm{SSS,PD}\overline{m}}$ and $A_{\mathrm{SSS,HSC}m}$ (cf. Equation 7).

In addition to the accuracies of the four models, the accuracies $A_{\mathrm{SSS,PD}\overline{m}}$ and $A_{\mathrm{SSS,HSC}m}$ are shown as gray lines in Figure 7. $A_{\mathrm{SSS,PD}\overline{m}}$ indicates the accuracy, that $\hat{f}_{\mathrm{SSS,PD}}$ decides right regarding samples where it is certain. If $r \to 0.5$ it converges to the accuracy $A_{\mathrm{SSS,PD}}$ as all decisions are made by the SSS with PD samples as input. If $r \to 0$ it converges to 100%. This makes sense, since only those samples are considered for which the model is 100% sure. However, it is not self-evident as a 100% certain model can also decide wrongly. $A_{\mathrm{SSS,HSC}m}$ indicates the accuracy of $\hat{f}_{\mathrm{SSS,HSC}}$ for the samples, where the SSS of the PD is uncertain. The accuracy of these samples given by $\hat{f}_{\mathrm{SSS,PD}}$ has to be below $A_{\mathrm{SSS,PD}}$ as $A_{\mathrm{SSS,PD}\overline{m}}$ always is above. The graph of $A_{\mathrm{SSS,HSC}m}$ shows that $\hat{f}_{\mathrm{SSS,HSC}}$ is already correct for over 93%, except for r near 0.5, of the samples for which $\hat{f}_{\mathrm{SSS,PD}}$ is uncertain and has an accuracy less than $A_{\mathrm{SSS,PD}} = 88.70\%$. If $r \to 0$, then $A_{\mathrm{SSS,HSC}m} \to A_{\mathrm{SSS,HSC}}$.

A_{CS} indicates the accuracy of the cascaded system. If $r \to 0.5$, the accuracy converges to the accuracy of the SSS with PD signals as input since that system

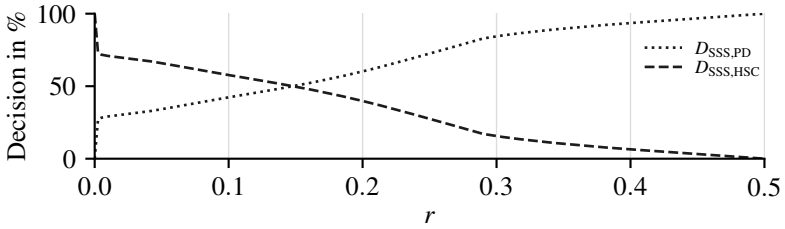

Figure 8: Final decisions. Shown are the percentages of the decisions in the cascaded system made by $\hat{f}_{\text{SSS,PD}}$ or $\hat{f}_{\text{SSS,HSC}}$ as a function of r.

makes all decisions; if $r \to 0$, A_{CS} converges to the accuracy of the SSS with HSC images as input, since then that system makes all decisions. With decreasing r, A_{CS} first follows $A_{\text{SSS,PD}\overline{m}}$ and then approaches $A_{\text{SSS,HSC}}$. Thereby, with $A_{\text{SSS,PD}\overline{m}}$ being the accuracy of samples using only the first classifier and $A_{\text{SSS,HSC}m}$ being the accuracy of the remaining samples resulting from the SSS of the HSC images, A_{CS} has to be in between of both (cf. Equation 7). At $r = 0.053$, A_{CS} has its maximum of 96.31%, which is greater than $A_{\text{SSS,HSC}}$.

Figure 8 shows what percentage of decisions in the CS are made by the respective system depending on r, where $r \to 0$ means that all decisions are made by the SSS with HSC images as input and $r \to 0.5$ that all decisions are made by the SSS with PD signals as input. For r near 0, an abrupt increase or decrease can be seen in the graphs. This is due to the distribution certainties of the SSS with the PD signals as input. This classifier is able to detect many anomalies with certainties near 100% as anomaly. For references, on the other hand, very few samples are classified with certainties near 100%. An explanation why the neural network learns in that way is the following: In the considered data some anomalies are clearly recognizable because they differ from the reference data. Hence, they are easily recognizable by the network. However, there are anomalies that look similar to reference data and are therefore very difficult to distinguish from references. Thus, the classifier does not often decide with certainties near 100% for a reference.

The number of parameters $P_{\text{SSS,PD}}$, $P_{\text{SSS,HSC}}$, P_{MSS} and P_{CS} in the inference of the systems, which is a rough indication for how long it takes the models to process data, are shown in Figure 9. Thereby, the MSS has the most parame-

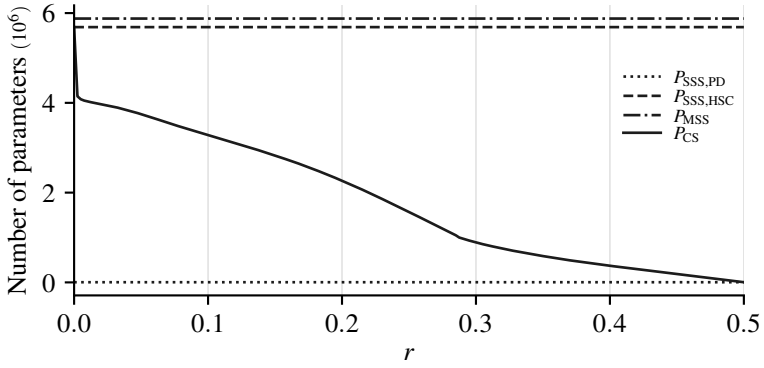

Figure 9: Number of parameters in the inference of the systems as measure for the computational effort. The concrete number of parameters are $P_{SSS,PD} = 3657$, $P_{SSS,HSC} = 5\,687\,233$ and $P_{MSS} = 5\,877\,961$.

ters, the SSS with the HSC images as input the second most and the SSS with the PD signals the least. The estimated computational effort of the CS results proportionally to how many samples are subjected to the first classifier only and how many are subjected to the first and second. If the decision is made by the SSS with the PD signals only, the estimated computational effort of the CS is equal to that of the SSS. With decreasing r the computational effort becomes larger. If $r \to 0$, than $P_{CS} \to P_{SSS,PD} + P_{SSS,HSC}$. If the CS has half of the final decisions made by the SSS with PD signals and the other half by the SSS with the HSC images, the estimated computational effort is reduced by 49.94% leading to an accuracy of 95.96%. Of interest is the region where the accuracy of the CS exceeds that of the SSS with the HSC images as input. At the point where A_{CS} is at its maximum and higher than $A_{SSS,HSC}$, the estimated computational effort is reduced by 34.46% compared to the SSS with HSC images.

6 General Cascaded System

The two-stage cascaded system proposed in this paper can be generalized. The structure of a general cascaded system (GCS) is shown in Figure 10. In contrast

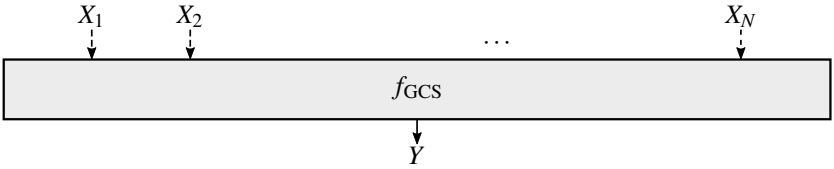

Figure 10: General cascaded system (GCS). The model f_{GCS} has $X_1, X_2, ..., X_N$ as inputs leading to a quality assessment Y.

to the two-staged cascaded system limited to two inputs, the GCS consists of N different inputs $X_1, X_2, ..., X_N$, which are (partially) processed together leading to a quality assessment Y. By that, the inputs do not necessarily have to come from different sensors, e.g. one sensor could provide several features as different inputs. Besides, possible inputs could also be metadata. The following structures of a cascaded system are possible, for example:

In contrast to the two-stage CS, the different input data $X_1, X_2, ..., X_N$ do not have to be divided into equal time periods. For example, considering the data used in this paper, input X_1 could make a decision for each entire weld seam. Further inputs can then work with smaller time periods in order to locate anomalies more precisely.

Unlike the proposed two-stage CS, which at uncertainty of a first classifier activates a second classifier, a GCS could exchange further information like learned features or certainties between classifiers. In doing so, the information flow does not have to be unidirectional. For example, speaking of a CS with two inputs, after making a first uncertain decision, the second classifier may return learned information to the first one in order to improve the first classifier's upcoming decision-making.

If there are several inputs, the system can decide which data $X_2, ..., X_N$ should be processed further depending on the results of the first input X_1. For example, if a first result supposes a specific anomaly among several, the next step evaluates exactly those inputs, that provide further information about that anomaly. The same principle can also be applied when choosing the initial inputs: Instead of using a fixed number of inputs in the first step, based on the decision of the previous step, initial inputs are selected. For example, if certain

anomalies are known to take a certain time, it could be useful to give attention to the inputs that can deal with the anomaly during that time period.

7 Summary and Outlook

Four different models for quality inspection of metal plates welds, namely two single-sensor system (SSS), a multi-sensor system (MSS) and a cascaded system (CS) have been considered. The CS presented in this paper is characterized by the fact that not all available data are analyzed to get the quality weld; only cleverly selected data are evaluated. The different models consisting of neural networks are compared in terms of their accuracy and estimated computational effort since fast but still precise quality inspection is needed in today's quality monitoring of welds. The data acquisition was done from scratch in the laboratory since no public data sets are available, whereat different anomalies such as spatters or gaps were inserted. Thereby, two common sensor methods for examining weld quality, namely a photodiode (PD) and a high-speed camera (HSC) have been used.

Among the four considered models, the SSS with PD signals as input has the fewest parameter but also the lowest accuracy. In contrast, the MSS has the highest accuracy but also the most parameters. If the CS has half of the final decisions made by the SSS with PD signals and the other half by the SSS with the HSC images, the estimated computational effort is reduced by almost 50% compared to the SSS with the HSC images as input. By that, the accuracy is only reduced by 0.25% from 96.21% to 95.96%.

Based on the CS, a generalized cascaded system (GCS) for quality inspection is proposed, which arbitrarily combines any number of sensors in order to get a quality assessment. Further work can deal with the possibilities that the GCS proposes, like the extension to more than two sensors or the incorporation of process knowledge. In addition to the used neural network architectures within the different models for feature extraction and classification, future work could include other architectures or methods based on classical machine learning. Furthermore, a more detailed analysis of the computational effort including other variables besides the number of parameters should be realized.

Moreover, the system can be extended to distinguish not only between anomaly and reference, but also between different anomalies.

References

[1] M. M. Atabaki, N. Yazdian, J. Ma, R. Kovacevic. “High power laser welding of think steel plates in a horizontal butt joint configuration”. In: *Optics & Laser Technology*, vol. 83, pp. 1–12. 2016.

[2] 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”. In: *Optics and Lasers in Engineering*, vol. 74, pp. 47–58. 2015.

[3] 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”. In: *Journal of Materials Processing Technology*, vol. 256, pp. 1–12. 2018.

[4] 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, pp. 1762–1767. 2001.

[5] A. G. Paleocrassas and J. F. Tu. “Inherent instability investigation for low speed laser welding of aluminum using a single-mode fiber laser”. In: *Journal of Materials Processing Technology* vol. 210, no. 10, pp. 1411–1418. 2010.

[6] A. Molino, M. Martina, F. Vacca, G. Masera, A. Terreno, G. Pasquettaz and G. D’Angelo. “FPGA implementation of time–frequency analysis algorithms for laser welding monitoring”. In: *Microprocessors and Microsystems* vol. 33., no. 3, pp. 179–190. 2009.

[7] S. S. Rodil, R. A. Gómez, José M. Bernández, F. Rodríguez, L. J. Miguel and José R. Perán. “Laser welding defects detection in automotive industry based on radiation and spectroscopical measurements”. In: *The International Journal of Advanced Manufacturing Technology* vol. 49., pp. 133–145. 2010.

[8] F. Kong, J. Ma, B. Carlson and R. Kovacevic. “Real-time monitoring of laser welding of galvanized high strength steel in lap joint configuration”. In: *Optics & Laser Technology*, vol. 44, pp. 2186–2196. 2012.

[9] 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”. In: *NDT & E International* vol. 42., no. 1, pp. 56–63. 2009.

[10] M. Thornton, L. Han and M. Shergold. “Progress in NDT of resistance spot welding of aluminium using ultrasonic C-scan”. In: *NDT & E International*, vol. 48, pp. 30–38. 2012.

[11] K. Wasmer, T. Le-Quang, B. Meylan, F. Vakili-Farahani, M. P. Olbinado, A. Rack and S. A. Shevchik. “Laser processing quality monitoring by combining acoustic emission and machine learning: a high-speed X-ray imaging approach”. In: *Procedia CIRP*, vol. 74, pp. 654–658. 2018.

[12] S. Shevchik, T. Le-Quang, B. Meylan, F. V. Farahani, M. P. Olbinado, A. Rack, G. Masinelli, C. Leinenbach and K. Wasmer. “Supervised deep learning for real-time quality monitoring of laser welding with X-ray radiographic guidance”. In: *Sci. Rep.*, vol. 10, no. 1, p. 3389. 2020.

[13] 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)*, pp. 1–10. 2021.

[14] S. Tsukamoto. “High speed imaging technique Part 2 – High speed imaging of power beam welding phenomena”. In: *Science and Technology of Welding and Joining*, vol. 16, no. 1, pp. 44–55. 2011.

[15] M. Jäger, S. Humbert and F. A. Hamprecht. “Sputter Tracking for the Automatic Monitoring of Industrial Laser-Welding Processes”. In: *IEEE Transactions on Industrial Electronics*, vol. 55, pp. 2177–2184. 2008.

[16] M. Jäger and F. A. Hamprecht. “Principal Component Imagery for the Quality Monitoring of Dynamic Laser Welding Processes”. In: *IEEE Transactions on Industrial Electronics*, vol. 56, pp. 1307–1313. 2009.

[17] 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)*, pp. 1377–1381. 2017.

[18] B. Zhou, T. Pychynski, M. Reischl, E. Kharlamov and R. Mikut “Machine learning with domain knowledge for predictive quality monitoring in resistance spot welding”. In: *Journal of Intelligent Manufacturing*, vol. 33, no. 4, pp. 1139–1163. 2022.

[19] 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”. In: *Journal of Manufacturing Processes*, vol. 77, pp. 54–62. 2022.

[20] T. S. Yun, K. J. Sim and H. J. Kim. “Support vector machine-based inspection of solder joints using circular illumination”. In: *Electronics Letters* vol. 36, no. 11, p. 1. 2000.

[21] X. Hongwei, Z. Xianmin, K. Yongcong and O. Gaofei. “Solder joint inspection method for chip component using improved AdaBoost and decision tree”. In: *IEEE Transactions on Components, Packaging and Manufacturing Technology* vol. 1, no. 12, pp. 2018-2027. 2011.

[22] H. Wu. “Solder joint defect classification based on ensemble learning”. In: *Soldering & Surface Mount technology*, vol. 29, no. 3, pp. 164–170. 2017.

[23] 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”. In: *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 3, no. 3, pp. 516–522. 2013.

[24] A. Krizhevsky, I. Sutskever and G. E. Hinton. “Imagenet Classification with Deep Convolutional Neural Networks”. In: *Communications of the ACM* vol. 60, no. 6, pp. 84–90. 2017.

[25] 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)*. pp. 2818–2826. 2016.

[26] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le and H. Adam. “Searching for MobileNetV3”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* pp. 1314–1324. 2017.

[27] Y. Yang, L. Pan, J. Ma, R. Yang, Y. Zhu, Y. Yang and L. Zhang. “A High-Performance Deep Learning Algorithm for the Automated Optical Inspection of Laser Welding”. In: *Applied Science* vol. 10, no. 3. 2020.

[28] P. Stritt, M. Boley, A. Heider, F. Fetzer, M. Jarwitz, D. Weller, R. Weber, P. Berger and T. Graf. “Comprehensive process monitoring for laser welding process optimization”. In: *Proc. SPIE 9741, High-Power Laser Materials Processing: Lasers, Beam Delivery, Diagnostics, and Applications V, 97410Q*. 2016.

[29] 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”. In: *Journal of Manufacturing Systems*, vol. 51, pp. 87–94. 2019.

[30] Y. Yang, R. Yang, L. Pan, J. Ma, Y. Zhu, T. Diao and L. Zhang “Real-time monitoring of high-power disk laser welding statuses based on deep learning framework”. In: *Journal of Intelligent Manufacturing*, vol. 31, pp. 799–814. 2020.

[31] T.-H. Kim, T.-H. Cho, Y. S. Moon, S. H. Park. “Visual inspection system for the classification of solder joints”. In: *Pattern Recognition*, vol. 32, no. 4, pp. 565–575. 1999.

[32] 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*, pp. 2440–2445. 2007.

[33] 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”. In: *Applied Soft Computing* vol. 102, no. 107115. 2021.

[34] Z. Wang, W. Yan and T. Oates. “Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline”. In: *International Joint Conference on Neural Networks (IJCNN)*. 2017.

[35] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam. “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”. In: *arXiv:1704.04861*. 2017.