Predictive Quality for Battery Contacting

1 Task Description

In the following task, measured process data during (inline) the laser beam welding process of electrical connections between battery cells and their connectors is provided. Two data sets were obtained using two different photodiode-based sensors which recorded the back reflected laser radiation during laser beam welding to detect process defects, such as spatters, pollution, or holes at the weld seam´s surface. For the labeling of the measured signals (also called time series) an automated algorithm was used, see Figure 1, and the signal was divided into five parts. For the sake of completeness, it is mentioned that the signal directly corresponds to the position of the weld seam (Figure 2).

The main two objectives are:

1. Determining possible correlations between the measurements of sensor 1 and 2
2. Classifying the weld seam in four categories (OK , not OK, pollution, out of threshold)

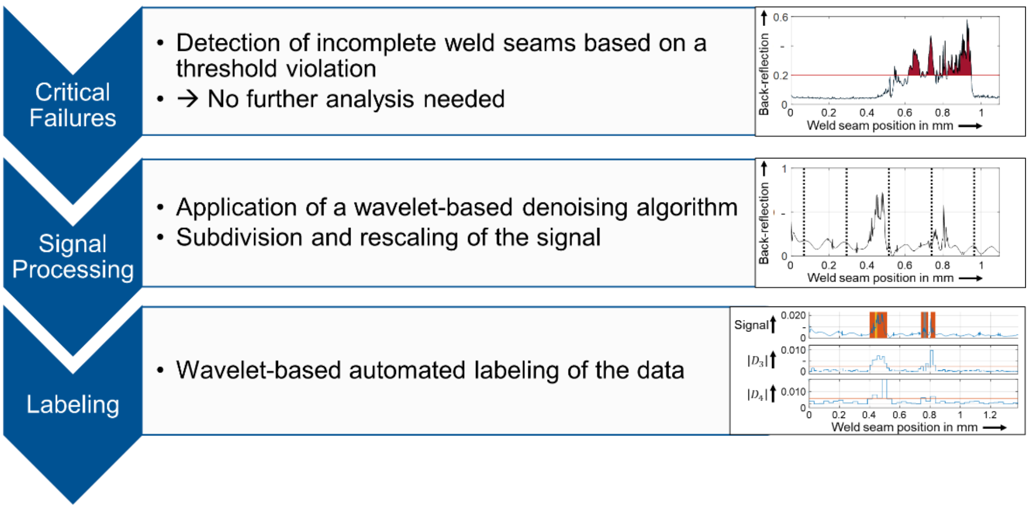

Figure 1: Functional principle of the automated labeling algorithm for the measured back reflected laser radiation

1.1 Process description and resulting datasets

The laser power was kept constant at a constant power during the experiments, whereas the feed rate was varied in a certain parameter range. A sensor with three photodiodes (sensor 1) and a sensor with one photodiode (sensor 2) was used for recording the process emissions during laser beam welding of pure copper stripes to a pure copper sheet (Figure 2).

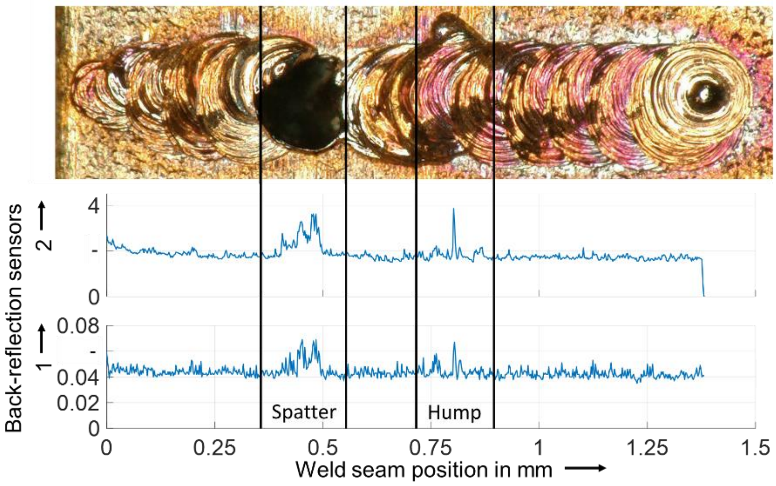

Figure 2: Experimental set up to manufacture the weld samples and to measure the process emissions (Weiss et al. 2022)

The sensor 1 covers different wavelength ranges included in the process emissions:

- plasma signal (PI_Raw)
- temperature signal (T_Raw)
- reflection signal (R_Raw)

Since the plasma and temperature signal do not contain significant information, only the reflection signal was saved in the data set.

The sensor 2 consists of a single photodiode and can only measure the reflection of the laser radiation. Sensor 1 and 2 show a linear response in their voltage values for an increasing intensity. The structure of the data set is summarized in Table 1.

Table 1: Structure of the data set

Column	Description
“not OK”	For a “1”, the weld seam is not OK.
“signal”	For a “1”, the signal exceeds a certain threshold. The threshold is the same for the entire data set.
WD40	For a “1”, the surface between the stripe and sheet was polluted with the lubricant WD40.
Gleitmo	For a “1”, the surface between the stripe and sheet was polluted with the lubricant Gleitmo.
LWMID_1	The LWMID_1 is a sequential number and is unique for each weld seam
LWMID_2	The LWMID_2 indicates a certain part of a weld seam (LWMID_1) and has a range from 1 (beginning) to 5 (end)
Signal1_1 – 112	Each column contains a normalized raw value of the signal for sensor 1.
Signal1_dn_1 – 112	Each column contains a normalized raw value of the signal with a noise reduction for sensor 1.
Signal2_1 – 112	Each column contains a normalized raw value of the signal for sensor 2.

1.2 Objective 1: Determination of possible correlations between the inline and online data sets

Based on the datasets, a possible correlation between the measured data has to be investigated for sensor 1 and 2. The correlations between the measurements can, for instance, be included in the contained frequencies. Certain peaks in the signal's frequencies can indicate anomalies correlated to weld seam defects caused by process instabilities.

1.3 Objective 2: Classifying the weld seam in four categories (penetration weld, humping, no connection, connection)

Due to the superimposition of many effects that occur highly dynamic and cannot be observed separately, the signals contain anomalies related to process instabilities and are included in different frequency spectrums. Based on the data sets, the weld seams shall be classified in the first group in the following three categories:

- OK
- Not OK
- Signal value exceeded

As a second group, the weld seams need to be classified to one of the following categories (only one of the two alternatives):

Alternative 1:

- Lubricant
- No lubricant

Alternative 2:

- WD40
- Gleitmo

For the training of a classification model, the signals as well as features (e.g. the frequency) can be used as input. Based on a test dataset, the trained classification model should be used to determine the weld seam quality belonging to one category of each of the two groups listed above.

1.4 Research questions

As the provided dataset allows a variety of analysis, some research questions are provided, which should guide the way to develop sensible machine learning models.

Week 1: Data exploration

What does the dataset look like? What kind of input features do you have (numeric, categorical, time series etc.)? Do you have to conduct the train/test-split manually? How many sensor signals do not change at all over time? Are different sensor signals correlated to each other? What do the labels look like?

Week 2: Data preparation and feature engineering

How can the measurement signal be suitably pre-processed? How can process instabilities be detected? Will you have to extract extra features or labels to answer the objectives 1 and 2? What features are relevant and suitable for the training of the model?

Week 3: Modeling

Which models can be used for the objectives? What are benchmark models for a classification tasks found in literature (have a look at Google Scholar, sciencedirect.com or ieeexplore.ieee.org)? After choosing a model, what are the critical parameters, hyper-parameters and how should you decide on them?

How can you deal with a small data set? Are there methods to extend the data set? Are there methods to artificially generate time series data?

Week 4: Results and model insights

How good does the model perform on the test data? What are levers to increase the model performance? Is hyper-parameter tuning possible and sensible? Is your model sensitive to random seeds and train/test-splits? How robust is the model when reducing the number of training observations? If you had more time, what would your next steps be? How applicable is your model in real-world use cases?

Week 5: Report

Can you motivate your research by a business case? What visualizations can communicate your results in a consumable way?

2 Literature

Grabmann et al. 2020

Grabmann, S.; Tomcic, L.; Zaeh, M. F.: Laser beam welding of copper foil stacks using a green high power disk laser. Procedia CIRP 94 (2020), S. 582-586.

Haddad et al. 2022

Haddad, E.; Chung, W. S.; Katz, O.; Helm, J.; Olowinsky, A.; Gillner, A.: Laser micro welding with fiber lasers for battery and fuel cell based electro-mobility. Journal of Advanced Joining Processes 5 (2022) 14, S. 100085.

Stadter et al. 2020

Stadter, C.; Kick, M. K.; Schmoeller, M.; Zaeh, M. F.: Correlation analysis between the beam propagation and the vapor capillary geometry by machine learning. Procedia CIRP 94 (2020) 2, S. 742-747.

Stadter et al. 2020

Stadter, C.; Schmoeller, M.; Rhein, L. von; Zaeh, M. F.: Real-time prediction of quality characteristics in laser beam welding using optical coherence tomography and machine learning. Journal of Laser Applications 32 (2020) 2, S. 22046.

Weiss et al. 2022

Weiss, T.; Kick, M. K.; Grabmann, S.; Geiger, C.; Mayr, L.; Wudy, K.; Zaeh, M. F.: holistic approach for an intelligent laser beam welding architecture using machine learning for the welding of metallic bipolar plate for polymer electrolyte membrane fuel cells. (Paper accepted)

3 Report

- One report per group in German or English
- Identify, which group member contributed to each chapter in which share (for example: Introduction: All 20 %, Methods: Person A 50 %, Person B 50 % …)
- 10 – 12 pages (for 5 group members)

Structure of the report:

1. Introduction (15 %)
   - Why is this topic dealt with? Can you motivate your topic by a business case?
2. Methods/Experiments (20 %)
   - How was this topic dealt with?
3. Results (20 %)
   - What were the outcome and results?
4. Discussion (20 %)
   - How can the results be interpreted, what are the consequences and limitations?
5. Summary (5 %)
   - What could be next steps and key learnings?
6. Documentation of the Code (20 %)

Introduction

- Introducing the reader to the topic
- The reader should understand the general topic and the motivation.

Methods/Experiments

- Description of the entire data-handling and analysis process (e.g. application of the KDD process)
- Description of the applied methods and models
- Descriptions should enable the reproduction off the results presented in the next chapter
- Why did you choose the individual methods and models?

Results

- Clear presentation of the results
- Precise and meaningful labeling of illustrations and diagrams

Discussion

- Core of the report
- Interpretation and evaluation of the results
- Direct references to the results of the previous section

Summary

- Indicate the key findings you had and the future research you would conduct if you had more time

Annex

- Additional plots
- Code

General remarks – Citations

- No copyright infringement, will be counted as fraud
- Indication in the text with first author and year, e.g. MUSTERMANN 2009
- Add references to illustrations from other authors

General remarks – Bibliography

- Required for all sources used
- Information in the bibliography with authors, title, journal, publisher, edition, year of publication and page number
- For online sources, the main page with date of the last access

General remarks – Language and expression

- Short and meaningful sentences
- Precise writing style
- Avoidance of unnecessary filler words
- No first person, no “one can/…” constructions
- Cover page without page number

General remarks – Content

- Methodological procedure
- Originality
- Complexity
- Integration of domain knowledge
- Interpretation of results and outlook
- …

4 Presentation

- 20 minutes presentation
- Q&A session – everyone should answer at least once.ga