Monitoring of Laser Welding via Coaxial Imaging and Acoustic Emission

Master of Science in Technology Thesis
University of Turku
Department of Computing
Robotics and Autonomous Systems
Smart Systems Laboratory
December 2025
Henrique Hiram Libutti Núñez

Supervisors:
Dr. Widhi Atman
Dr. Kandice Ribeiro
Prof. Wallace Bessa

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

UNIVERSITY OF TURKU
Department of Computing

Henrique Hiram Libutti Núñez: Monitoring of Laser Welding via Coaxial Imaging and Acoustic Emission

Master of Science in Technology Thesis, 51 p.
Robotics and Autonomous Systems
Smart Systems Laboratory
December 2025

Laser Welding (LW) offers multiple advantages compared to previous welding technologies and is a key technology in battery manufacturing and aerospace industry. However, process discontinuities hinder broader market adoption for use cases such as welding thick plates for structural steel. Quality inspection systems and automated process control are a hot topic that promise to increase the automation level in LW and deliver higher quality manufactured components. The aforementioned automated systems require computer software that measures process parameters from sensors, and in this context previous studies demonstrated the potential of coaxial imaging and Acoustic Emission (AE) for LW monitoring, of which the main advantages are the online process monitoring capability and ease of adaptation into existing LW systems. Specifically, laser-illuminated cameras offer great advantages for monitoring since they use band pass filters, removing process noise and delivering high quality images; In parallel, the development of optical microphones opened vast possibilities for acoustic monitoring of signals in the order of MHz. This thesis studies two different Machine Learning (ML) methods for predicting process parameters in LW. In the first study, coaxial images were used to predict the penetration depth using supervised learning and Optical Coherence Tomography as ground truth data. In the second study, AE was used for laser power prediction, where two alternatives were compared: a Deep Learning model and a statistical feature extraction library combined with out-of-the-shelf ML regressor. For evaluating both alternatives, train-test validation was used, as well as regression scoring functions. The Coefficient of Determination (R^2) was applied for both methodologies. For image-based prediction of the keyhole depth, the R^2 was of 0.59 for the best strategy, and for the AE-based laser power prediction, a R^2 of 0.93 was obtained in the best strategy. Both results are successful indicators of using coaxial imaging and optically sourced AE with ML to predict LW parameters. In the case of imaging, improvements are suggested to obtain better results, while in the case of AE, the results strongly indicate the success in capturing the data variance. In a wider perspective, this work presents advancements in process monitoring for LW that could be applied in quality control, self-correcting systems, and parameter optimization.

Keywords: process monitoring, laser welding, machine learning, acoustic emission, coaxial imaging

Contents

List of Figures

List of Tables

1 Introduction

The 21st century demands technological solutions to address climate change and labor shortages. Industry 4.0, characterized by digitalization and automation of manufacturing processes, promises more efficient production, safer working environments, and reduced waste [1, 2], enabling better energy usage and increased competitiveness. The shipbuilding industry, in particular, must adapt to a new scenario: maritime shipping alone contributes approximately 3 % of global greenhouse gas emissions, a figure expected to rise as trade activities intensify [3, 4]. In response to regulations from the International Maritime Organization and the European Union, alongside economic pressures, there is an ongoing effort to transition shipbuilding practices towards low emission designs [5–7]. Lower emission designs can be achieved by using higher-strength steels, which enable the construction of thinner and consequently lighter structures [8].

In modern shipbuilding, welding is one of the most critical manufacturing steps [9]. The industry is increasingly focused on welding thinner plates with minimal distortion, making Laser Welding (LW) a key enabling technology. LW operates with highly focused heat input, which creates a small Heat Affected Zone (HAZ) and allows for the high-speed joining of thin materials while preserving their structural integrity [10, 11].

Compared to traditional joining methods, LW offers superior precision and reduced thermal distortion, making it valuable for industries requiring high-quality

joints, such as automotive manufacturing [12], aerospace [13], and shipbuilding [9]. These characteristics not only enable the use of new materials but also have the potential to reduce overall production costs due to distortion rework.

1.1 Motivation

Despite its advantages, LW faces significant challenges that limit its broader adoption: porosity and cracking defects remain persistent issues that hinder certification and scalability across applications; furthermore, variations in process parameters and part preparation are additional factors for process instabilities [14]. To address these challenges, the scientific and engineering communities are developing autonomous manufacturing systems, capable of acquiring and interpreting process data used by self-correcting control mechanisms, with the goal of increasing process quality and reducing costs and defects [15, 16]. Current research increasingly adopts in-process sensing using coaxial cameras [17], microphones [18], pyrometers [19] and Optical Coherence Tomography (OCT) [20] combined with advanced data processing techniques such as statistical analysis and Deep Learning (DL). Autonomous manufacturing systems that can sense process conditions and adapt in real-time [21] offer significant benefits including streamlined production, cost reduction, material savings, and improved operator safety by removing humans from hazardous and unsafe environments.

1.2 Problem Statement and Research Objective

This thesis addresses the industrial need for robust, in-process sensing algorithms for LW. The research was guided by two primary questions:

- Can images from coaxial cameras be exploited for real-time keyhole depth prediction as a cost-effective alternative to OCT?

- Can laser power be predicted by high-frequency acoustic emission process signatures?

Therefore, the objective of this work is to evaluate new methodologies for LW monitoring. To fulfill this objective, two distinct methodologies have been proposed and validated: (i) A supervised learning model using coaxial images to predict keyhole depth, benchmarked against OCT measurements, and (ii) Signal processing pipelines for predicting laser power from acoustic emission data.

The research conducted to meet these objectives resulted in two peer-reviewed conference publications: 'In-situ monitoring and online prediction of keyhole depth in laser welding by coaxial imaging,' presented at the 14th CIRP Conference on Photonic Technologies (LANE 2024) [22], and 'Evaluation of Acoustic Emission as a Predictor of Laser Power in Laser Welding,' presented at the 20th Nordic Laser Materials Processing Conference (NOLAMP20) [23].

These contributions were developed within the 'CaNeLis' (Carbon-neutral lightweight ship structures using advanced design, production, and life-cycle services) project, a collaboration between academic and industrial partners in Finland. This research is part of the larger NEcOLEAP ecosystem, which aims to support the development of a carbon-neutral cruise ship by 2030.

Lastly, the data processing algorithms used in this work and the two aforementioned publications are available at the 'GitLab' code sharing platform under the repositories: (i) gitlab.com/henriquenunez/keyhole-depth-pred-img-lane2024; and (ii) gitlab.com/henriquenunez/laser-power-pred-optical-mic-nolamp20, both licensed under the EUPL.

1.3 Thesis Organization

This Master's Thesis is organized in seven chapters. The first and current chapter presents the introduction, motivation and relevance of this work. Chapter 2 contains the literature review for the development of this work, starting from work on the influence of process parameters on defects and how feedback control systems can improve the process. Chapter 3 details the experiments and algorithms for predicting the keyhole depth based on coaxial images. Chapter 4 details the experiments and algorithms for predicting the laser power based on optical microphone monitoring. Chapter 5 presents the results obtained by both methodologies. Chapter 6 draws conclusions based on the results from chapter 5. Finally, chapter 7 presents next steps that can extend and improve this research.

2 Laser Welding and Process Monitoring

Since the development of the first working laser by Maiman in 1960 [24], a whole new range of applications has been developed: lasers today are used in medicine, manufacturing, computing, research and more [25]. In the metal manufacturing industry specifically, lasers are typically used as a heat source to change the material from a solid to a liquid state. The transformative potential of this technology has been realized through its application in cutting, additive manufacturing, surface engineering, and welding [26, 27].

LW offers numerous advantages over traditional welding processes: by using a concentrated heat source, it minimizes both spatial distortion and the HAZ, while also allowing for higher processing speeds [27, 28]. Its high energy density and operational flexibility have led to widespread adoption in the fabrication of large-scale metallic structures, becoming a key process in areas such as automotive manufacturing [29, 30], aerospace [31], and shipbuilding [32].

LW can be performed in a manual setup, where an operator handles the laser delivery head, or in a robot-controlled setup, where a robotic arm performs the operation. Compared to manual operation, the robot-based approach offers greater repeatability and a larger working volume, whereas manual operation allows for more operational flexibility. Robot-controlled systems are often preferred in industrial

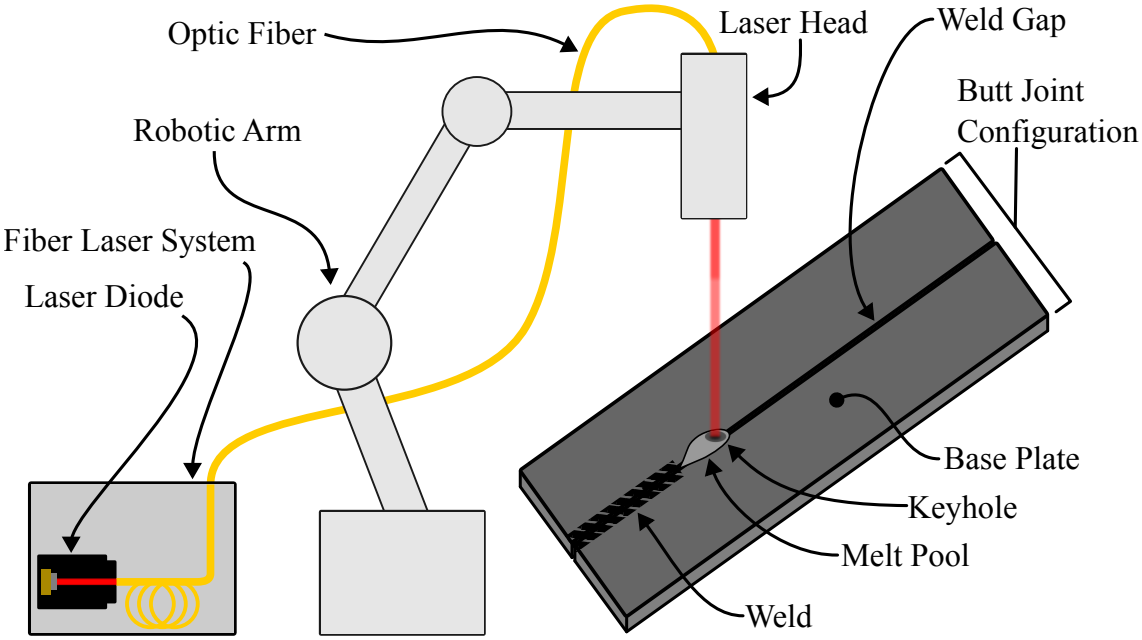

Figure 2.1: Scheme of robotic arm laser welding with material plates in a butt-joint configuration.

contexts due to higher operator safety and process repeatability [33]. Additional benefits of robotic LW include computerized parameter control (e.g., speed and laser power) and the ability to integrate sensors onto the welding head for in-process monitoring [34]. Figure 2.1 depicts a butt-joint LW setup controlled by a robotic arm.

2.1 Laser Welding Parameters and Defects

In LW, the main process parameters are laser power, welding speed, and focal distance [35]. Additional control can be achieved through beam shaping, which alters the power distribution pattern [36], and the use of shielding gas to create a protective atmosphere around the work region and prevent oxidation [28]. The material type and plate thickness ultimately dictate how these parameters influence weld quality [35]. Thicker materials, for instance, require a higher power input to compensate for the larger volume and increased thermal conduction.

Two main modes of LW exist: conduction mode and keyhole mode. The primary difference lies in the power density and the resulting formation of a keyhole [28, 30, 37, 38], which is a temporary cavity formed by material vaporization [39, 40]. Keyhole welding enables the joining of thicker plates, while conduction mode is used to minimize porosity and the HAZ in high-precision applications. The keyhole itself is subject to complex interaction of forces, including gravity, shielding gas pressure, and vapor plume dynamics [39, 41]. Table 2.1 provides a comparison of both LW modes, and Figure 2.2[1] illustrates the keyhole and associated phenomena.

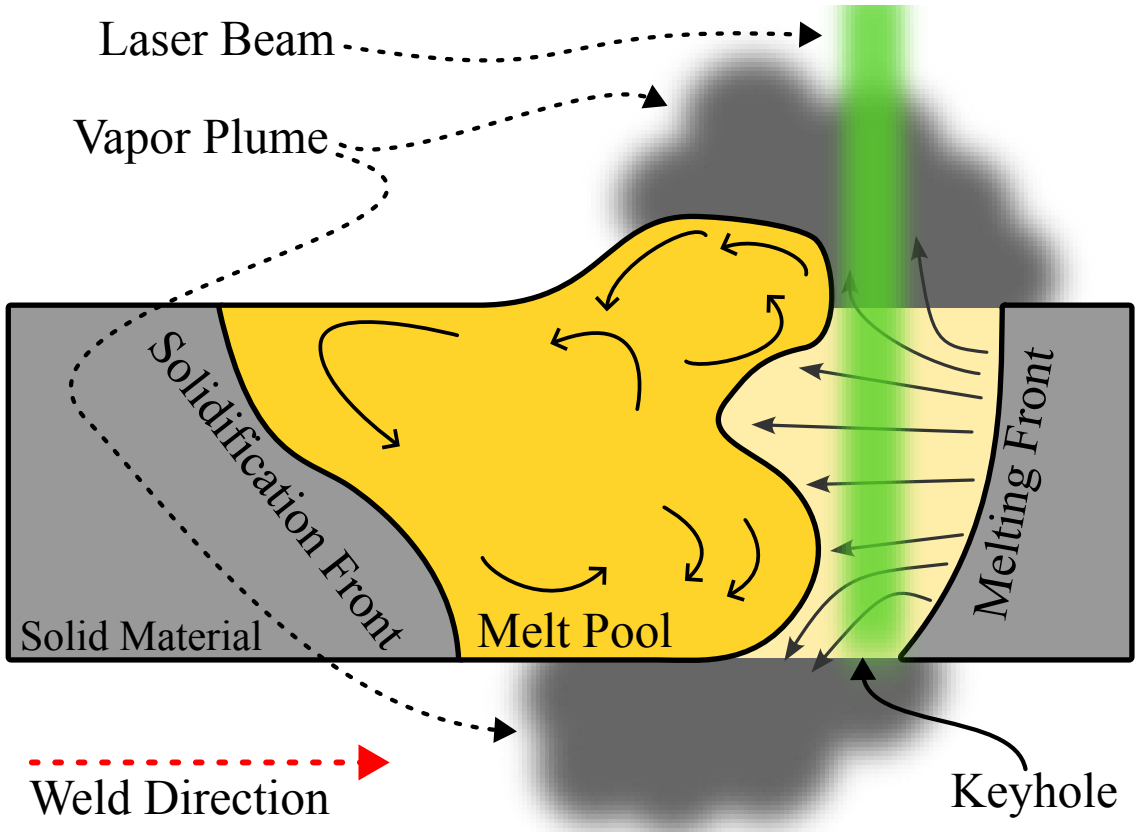

Figure 2.2: Scheme of keyhole in LW indicating associated phenomena. Adapted from [39].

A crucial phenomenon during keyhole welding is the formation of porosity, which

[1]Reprinted from Journal of Physics D: Applied Physics, Vol 46, Mickael Courtois, Muriel Carin, Philippe Le Masson, Sadok Gaied1 and Mikhaël Balabane, "A new approach to compute multi-reflections of laser beam in a keyhole for heat transfer and fluid flow modelling in laser welding"", Copyright 2013, with permission from IOP Publishing.

Table 2.1: Comparison between conduction mode and keyhole mode LW.

Mode	**Power Density**	**Penetration**	**Applications**	**References**
Conduction	Low	Shallow	Precision Aerospace EV Batteries	[30, 37]
Keyhole	High	Deep	Structures Shipbuilding	[28, 30]

occurs when parts of the keyhole become trapped during solidification [42]. These pores weaken the mechanical strength and fatigue resistance of the weld. Studies have investigated porosity formation in LW using X-ray imaging and Acoustic Emission (AE) monitoring, and have proposed strategies such as laser remelting for its removal [43]. Figure 2.3 depicts synchrotron X-ray images of pore formation in laser-welded parts.

Mitigating these issues often begins with parameter optimization. By selecting parameters that ensure complete fusion and uniform heat distribution, the overall weld quality can be improved to some extent [44]. However, defects may still appear due to process instabilities. A significant challenge in LW is optimization process itself, which involves testing numerous combinations of parameters and analyzing outcomes such as penetration depth, hardness, and the presence of defects like undercut or lack of fusion [45]. Traditionally, weld cross-sections are used to analyze penetration and characterize the HAZ, which is a time-consuming and destructive testing method [35].

To address these limitations, online process monitoring using head-mounted sensors has become an active area of research [46]. Process monitoring integrates sensors, domain-specific knowledge, and data analysis to create an automated system able to detect discontinuities during fabrication [40]. A broad variety of sensors exist in this regard, each responding to different physical inputs: process luminance, acoustic emissions, temperature, among others, which provide complementary in-

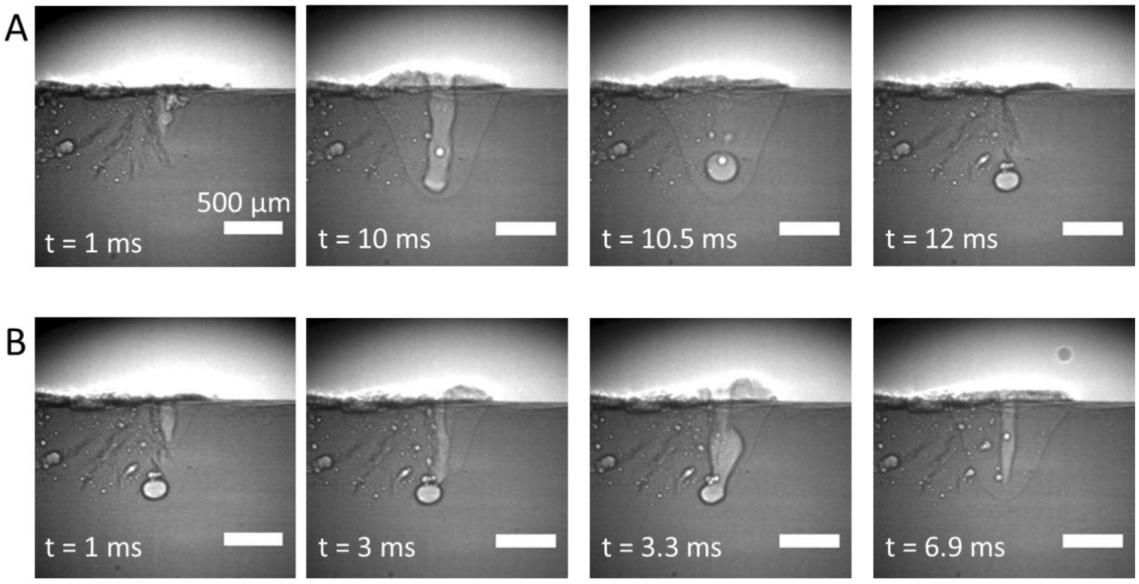

Figure 2.3: X-ray images of pore formation in LW (A) and pore removal via subsequent laser remelting (B). Originally published in [43]. Licensed under CC BY 4.0.

formation about the process state [40].

2.2 Process Monitoring

Monitoring keyhole behavior during LW is critical due to the complex thermo-fluid dynamics within the weld pool [47]. Insufficient penetration, for example, can lead to severe structural defects that compromise mechanical integrity and necessitate costly rework [48]. Therefore, process monitoring as a field has the goal of developing methods and systems capable of extracting useful information from the process. This information can afterwards be applied for quality control systems (with direct applications in part inspection and parameter optimization) and ideally in a self-controlling, automated LW system, with minimal human intervention and need of rework.

Among the various sensing modalities investigated for LW monitoring, coaxial cameras have emerged as a promising option due to their non-intrusive nature and compatibility with existing welding optics [14, 49]. Recent advancements in artificial

intelligence (AI), particularly DL, have significantly enhanced the ability to extract meaningful features from visual data [15]. This convergence of high-performance computing and sophisticated DL models presents new opportunities for real-time process monitoring.

In-process optical imaging uses cameras to capture real-time visual signatures of the weld pool and keyhole. While synchrotron X-ray imaging can reveal nuanced dynamics, such as pore formation and keyhole collapse with micrometer resolution [41], its use is restricted to laboratory settings due to its high cost and the hazardous environment created by ionizing radiation.

Alternatives to this type of imaging are cameras that use inherent process information, such as thermal radiation or laser back reflection from the workpiece; however those systems are unable to deliver satisfactory structure information (for example melt pool geometry) [50], which is required for a holistic process monitoring. Thermal cameras have been employed to analyze thermal distributions in laser material processing; such cameras however commonly fall into the problem of emissivity calibration which is often a major problem, and only few alternatives are commercially available as of recently [38].

A practical alternative to overcome those limitations is to employ laser-illuminated cameras: this methodology consists of a camera coupled with a narrow band pass filter, with a laser-diode simultaneously directed to the workpiece [50, 51]. In this system, external illumination reaches the workpiece, the reflected signal enters the optics, and the band-pass filter discards all but the spectrum of interest, delivering an image of the melt pool without noise from the process laser or light from the thermal emissions.

Previous studies have demonstrated the efficacy of coaxial imaging in related laser-based processes. For instance, [52] employed melt pool images to predict focal distances in directed energy deposition, while [53] utilized similar imaging techniques

to estimate final clad widths. These applications underscore the potential of visual monitoring in laser-based manufacturing.

Two-dimensional measurement systems are often insufficient for a thorough characterization of the process, as depth information can only be inferred. Inspection systems that can directly measure distance and obtain point cloud data offer a more comprehensive data source. Laser line scanning is one such methodology, employing laser triangulation to calculate the distance between the sensor and the workpiece with micrometer-level resolution [54]. A previous study by Sun et al. merged consecutive line scans into a point cloud and applied DL for defect detection [55]. However, despite their accuracy, it is challenging to apply line scanners in-process due to harsh environmental conditions [56].

A different 3D sensing strategy is OCT, which computes laser interference patterns to measure distance. In this technique, laser beams are directed onto the process area, and the returned beams are measured to determine the relative distance to the workpiece based on their interference pattern [49, 57, 58]. OCT can be used in a scanning mode to obtain a point cloud in real-time [59]. This methodology requires calibration and involves high equipment costs. However, OCT has several advantages: it can share optics with the main process laser, allowing for a coaxial setup, and it enables live process monitoring. This provides an online data source that can be used for downstream process control and optimization [58–60]. Figure 2.4[2] describes an OCT system used to measure the keyhole in LW. The illustration shows the optical integration of both laser sources, with the OCT measurement laser reaching the keyhole.

Offering a different perspective, Focus Variation Microscopy (FVM) is a post-process inspection method that performs 3D surface reconstruction. Previous work

[2]Reprinted from Optics and Lasers in Engineering, Vol 119, Meiko Boley, Florian Fetzer, Rudolf Weber, Thomas Graf, “A new approach to compute multi-reflections of laser beam in a keyhole for heat transfer and fluid flow modelling in laser welding"", Pages 56–64. Copyright 2019, with permission from Elsevier.

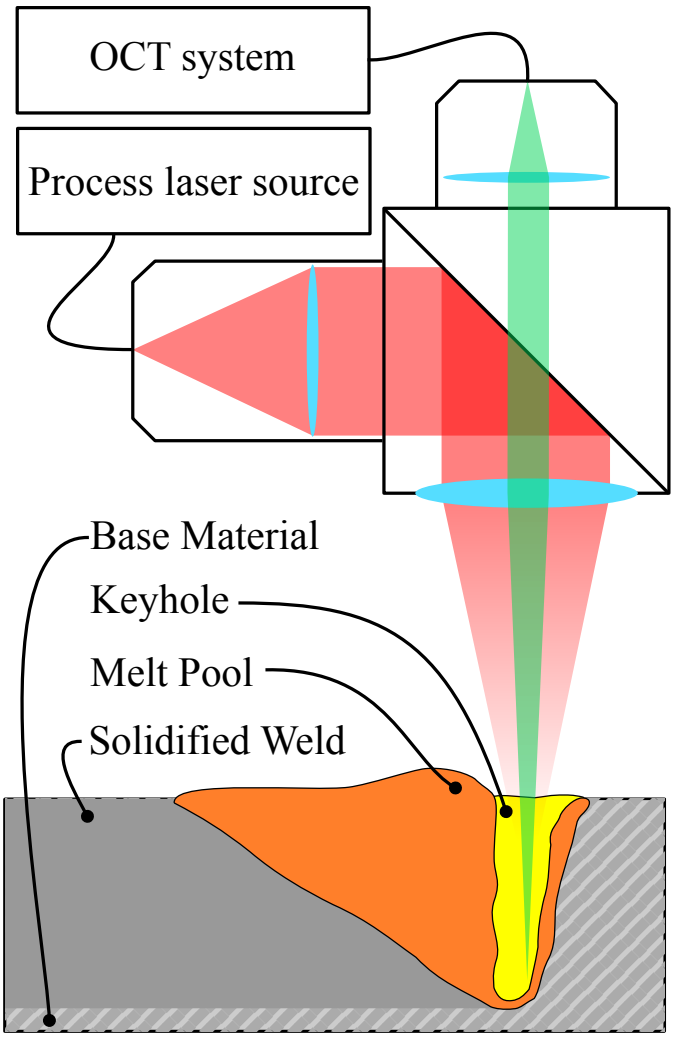

Figure 2.4: Scheme of an OCT setup for keyhole measurement in LW. Adapted from [57]

has demonstrated algorithms using FVM data to detect and characterize defects in friction stir welding [61], and other studies have shown that FVM can also be applied to LW [59].

In addition to optical analysis, AE sensing is another viable monitoring approach: this method is based on capturing the airborne and structure-borne vibrations generated by physical interactions during the process [62]. AE monitoring has been used to identify phenomena such as solidification cracking [63] and to estimate penetration depth in LW [64]. Conventional AE sensors are based on the piezoelectric principle, where vibrations in a crystal generate an electric signal; while inexpensive, these sensors typically have a limited operational frequency range of up to 150 kHz [63]. Since 2016, optical microphones based on laser interferometry have become commercially available [65]. The lack of moving parts, combined with high acquisition rates, makes optical microphones an excellent choice for process monitoring, with some systems offering sampling frequencies up to 4 MHz. Optical microphone monitoring is a topic of interest in laser material processing: one study found that optic-based AE provided enough information for calculating the laser ablation volume, focal position and material transitions [62], while another study [63] was able to correlate the laser signal to the Keyhole Depth (KD) in LW. Although the current cost is elevated, a broader market adoption is expected to reduce the cost of this equipment and lead to further innovations in equipment size and setups.

Other process monitoring strategies in LW include the use of pyrometers [19, 46], which measure temperature based on the material's thermal radiation and provide a direct, physically interpretable measurement, albeit with limited spatial resolution [38].

Taking into account the different signal acquisition methodologies, Table 2.2 summarizes the monitoring methods discussed. This work focuses on LW monitoring via camera, microphone, and OCT due to equipment availability and prior evidence

Table 2.2: Comparison of monitoring sensors in LW.

Method	**Advantages**	**Disadvantages**	**References**
Camera	Low cost, Algorithm availability	No direct depth information	[38, 51]
OCT	High resolution, Coaxial integration, Depth information	High cost, Requires calibration	[59, 60]
AE via Optical microphone	High acquisition rate, High signal-to-noise ratio	High cost	[62, 63]
Pyrometer	Direct temperature measurement	Requires emissivity calculation, Low spatial resolution	[19]
Line Scanner	High resolution Depth information	Normally post process	[54, 56]
FVM	High resolution Depth information	Post process only, Time-consuming	[59, 61]

of their success as data sources for monitoring systems.

2.3 Data Processing and Machine Learning

Continuous human monitoring of high-speed manufacturing processes is not economically viable, safe, or operationally reliable due to inherent limitations in sustained attention [66], therefore, process monitoring must operate in tandem with computational data analysis. Automated inspection and control systems require data processing methodologies able to transform raw sensor data into actionable information, which can then be used to modify process parameters and prevent defects [67].

Traditionally, data processing has been implemented through handcrafted algorithms based on expert knowledge and interpretable features [68, 69]. An example of handcrafted image processing is displayed in Figure 2.5, where the image is first transformed into grayscale (part A), next filtered using Gaussian blur (part B), and then subsequently transformed into a binary edge detection image (part C). In this image, it is possible to observe that this implementation is sensitive to inherent factors, such as noise from the environment and differences in light conditions. While effective for well-characterized processes, handcrafted features face challenges with newer data formats, previously unseen process variations, and novel sensor modalities [70, 71].

Image processing has historically received significant attention due to the ubiquity of cameras and the high information density of visual data. Advancements in camera hardware have improved spatial resolution (more pixels), temporal resolution (higher frame rates), and signal-to-noise ratio. In laser process monitoring, imaging systems have been used to measure melt pool area, analyze welding defects, and estimate bead width in directed energy deposition, among other applications [46].

In contrast to image-based monitoring, AE signals measure vibrations and offer a different physical interpretation of the process, outputting a one-dimensional time-

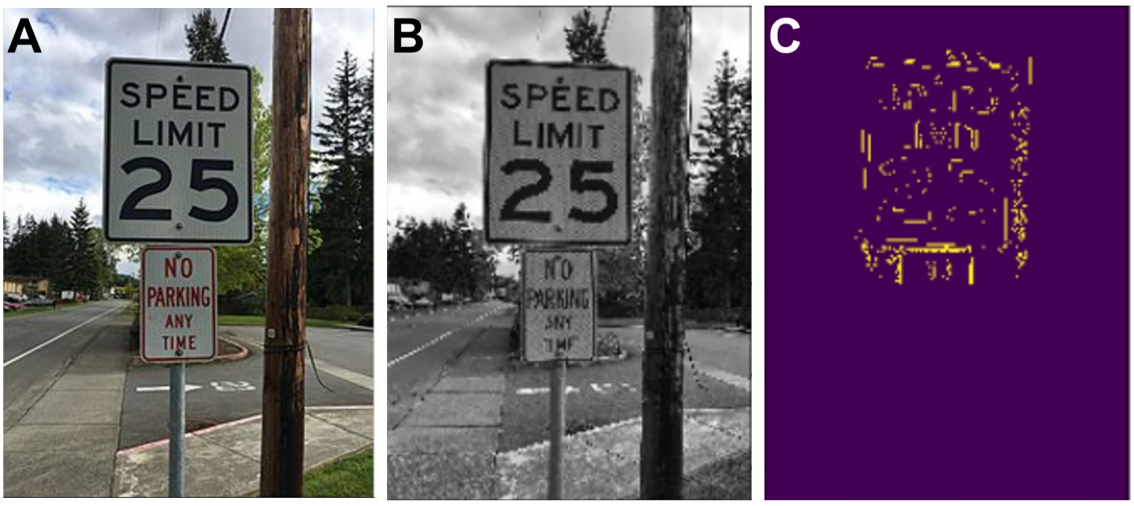

Figure 2.5: Example of handcrafted image processing steps. Adapted from the Wikipedia user Vidyakv. Licensed under CC-BY-SA 4.0.

dependent signal per channel. Common analysis techniques include time-domain analysis (e.g. root-mean-square, amplitude) and frequency-domain analysis such as the Discrete Fourier Transform or Wavelet decomposition [72]. Frequency-domain analysis allows for inspection of specific frequency components and has been employed for defect classification and process stability monitoring [64].

Machine learning (ML) has emerged as a powerful approach for process monitoring, particularly when the relationships between sensor inputs and process outcomes are too complex to model analytically. ML approaches can be categorized into supervised learning and unsupervised learning, which differ in their training methodologies. Supervised learning tunes model parameters to map inputs to known outputs, whereas unsupervised learning discovers underlying patterns using only input data. Classical ML methods, preceding neural network-based approaches, offer the advantage of requiring less training data.

Classical ML strategies include the Support Vector Regressor (SVR) and decision-tree-based methods such as Random Forest Regressor (RFR) and Gradient Boosting (GBR). SVR uses hyperplane separation in high-dimensional feature spaces to perform classification and regression tasks, and among its advantages are flexibility in

high-dimensional spaces and the use of kernels (e.g., radial basis function) for non-linear space separation [73]. Both RFR and GBR use a different strategy compared to SVR, based on decision-tree ensembles: they consist of a series of decision trees, and the aggregation of each tree result is the output of the whole ensemble [74]. A decision tree consists of decision nodes where the input features are used as a decision for further downstream decisions, successively reaching an outcome, which is the model result [74, 75]. Classical ML methods have been widely applied in manufacturing monitoring in previous studies [33, 45, 64, 71, 76].

While classical ML methods have been proven valuable, recent advances in DL have expanded the capabilities of automated process monitoring. DL models, particularly Convolutional Neural Networks (CNN), have become popular following landmark successes in computer vision. An architecture consisting of fundamental CNN components is depicted in Figure 2.6, demonstrating the data dimensionality transformation between layers the application of max-pooling and convolution. After the blocks of dimensionality reduction, the feature maps undergo flattening (represented by the 24 arrays of 16x16 elements) and this one-dimensional vector is applied through densely connected layers, transforming it into an array of 256 elements, and next into 128 elements. In this Figure, the input and output layers are not represented; however, the input is typically configured as a single channel image (e.g. 128x128x1), while the output can be set as a single floating point value or an array of predictions. Important architectural innovations include dropout layers [77], residual networks (ResNet) [78], VGG networks [79], U-Net architectures for [80], Vision Transformers [81] and variational autoencoders [82]. These architectures enable automatic feature extraction from raw data, eliminating the need for manual feature engineering.

However, despite advancements in data processing techniques, data scarcity remains a common challenge in niche application domains such as process monitoring,

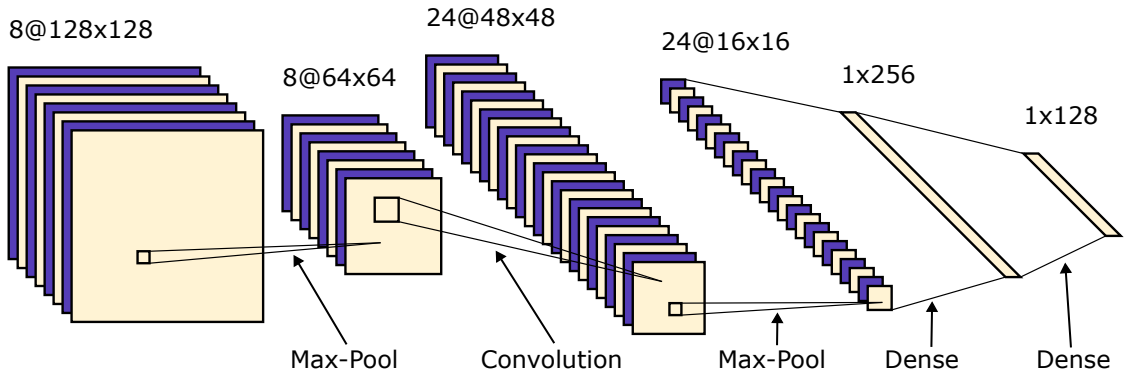

Figure 2.6: Depiction of a CNN architecture, consisting of the foundational blocks of Max-Pooling, Convolution and densely connected layers.

where obtaining millions of labeled examples is impractical [83]. This problem becomes more apparent when complex models, which require more internal parameters to achieve the higher complexity, also require larger amounts of training data to converge. Transfer learning addresses this issue by enabling the use of pre-trained models as feature extractors, allowing systems to benefit from the complex representations learned on large datasets without requiring extensive domain-specific data [83]. Figure 2.7 depicts a graphical representation of the transfer learning process, which starts with a model trained on a specific domain (e.g. image classification for landscape photography), in the upper part of the Figure, and is succeeded by a new application (e.g. melt pool image analysis). The new model reuses parts of the previous pre-trained model (usually the initial, less abstract layers) with a subsequent model that uses the extracted feature maps to predict the variable in question. The main advantage of transfer learning is the reuse of trained layers (which benefited from the generalization potential of large datasets). Figure 2.7 illustrates the transfer learning process; in which it is possible to observe that the layers of the CNN, in the upper half of the figure, are used in the new model, in the lower half of the figure. The custom model therefore establishes a bridge between the extracted features and the new output. This process has the consequence of requiring less training data [84].

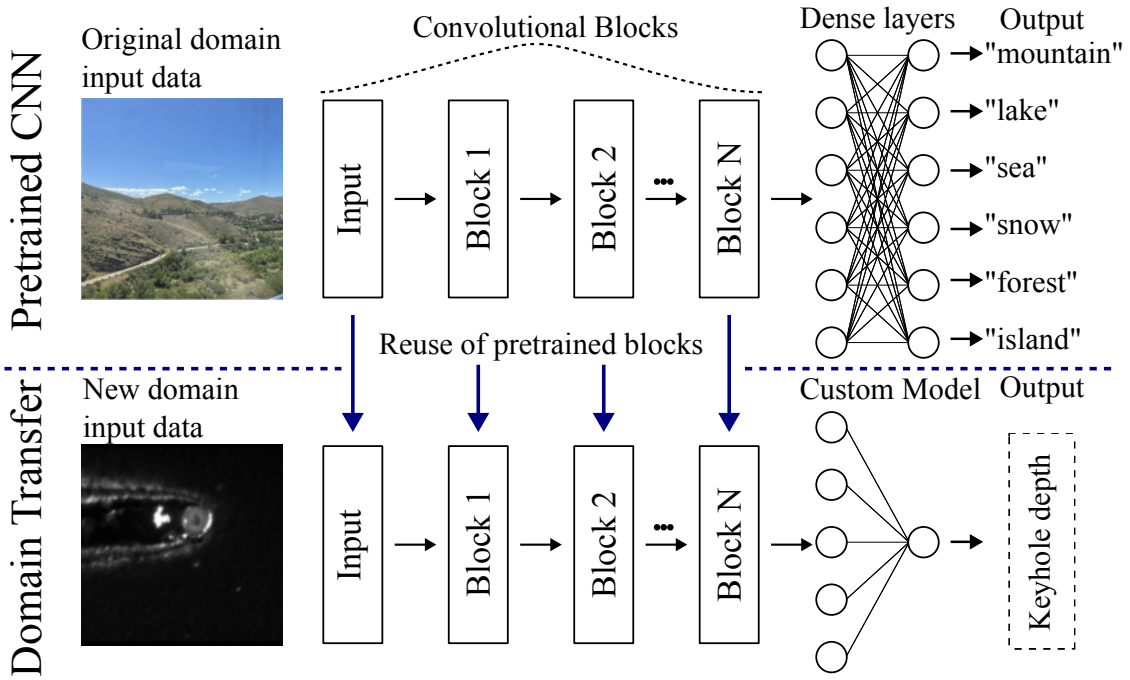

Figure 2.7: Schematic of the transfer learning process by using a pre-trained CNN for a new application domain.

2.3.1 Model Performance

In supervised ML, the dataset consists of pairs of input data and true values. This correspondence is used to train the model, such as the internal parameters of models (e.g. the weights of neural networks or the hyperplane normal vector in SVR), and can also be used to compare the model prediction of a certain input data with the ground-truth value for this data.

It is essential to quantify model performance in ML development, especially to establish comparisons between approaches and to evaluate generalization and predictive power. The evaluation techniques are directly dependent on the predictive task (i.e. regression versus classification). In this work, which focuses on regression tasks, a set of four metrics is used to evaluate regression performance: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Coefficient of Determination (R^2), and Root Mean Squared Error (RMSE). The mathematical formulations are presented as follows

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2.1}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2.2}$$

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{2.3}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2.4}$$

where y_i represents ground truth values, $\hat{y}_i$ represents predictions, and $\bar{y}$ is the mean of the observed values. MAE computes the average magnitude of prediction errors; MAPE normalizes MAE by the observed values; R^2 quantifies the proportion of variance in the data explained by the model; and RMSE calculates the standard deviation of the prediction errors. Therefore, a lower MAE, MAPE, and RMSE indicate better performance, while R^2 closer to 1.0 indicates better fit. This combination enables richer evaluation of the model in question.

Complementing the formulations in Equations 2.1 to 2.4 is dataset splitting into training and testing partitions: the data used in training and testing the models should be explicitly disjoint; otherwise, test information available during training compromises the validation of a ML model.

Common approaches of dividing a labeled dataset are the fixed train-test-split and the k-fold cross-validation (k-CV). Train-test splitting divides the dataset into two disjoint parts based on a ratio (i.e. 80:20), which can be done with or without randomization. In this example, the training set would have 80 % of the original dataset size, and the test set 20 %. This technique, however, fixes the train and test portions, and through a single partitioning it is not possible to measure the model

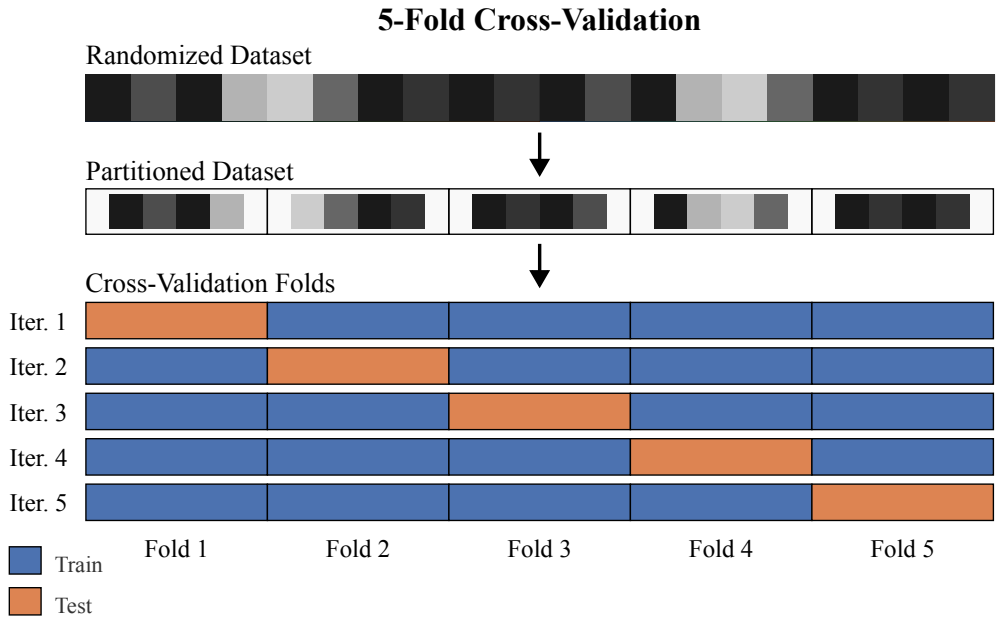

Figure 2.8: Scheme of k-fold cross validation.

performance across the entire dataset. An alternative to this limitation is the k-CV, which repeatedly partitions the data into k disjoint folds and iteratively trains on $k-1$ folds while evaluating on the held-out fold. This procedure is repeated k times, with each fold serving as the test set exactly once. As a result, k-CV provides k independent performance estimates, yielding more robust generalization metrics through averaging. Figure 2.8 illustrates the partitioning scheme for $k = 5$ folds.

3 Keyhole Depth Prediction Using Coaxial Camera Imaging

This chapter presents a supervised learning methodology for predicting KD in LW using coaxial camera imaging. The proposed approach evaluates whether KD can be monitored without expensive OCT sensors, thereby reducing system complexity and costs. The methodology combines pre-trained CNNs for feature extraction with supervised regression models to estimate KD from melt pool images. This approach could significantly simplify in-situ monitoring while maintaining or improving current quality assurance standards.

3.1 Methodology and Experimental Setup

The proposed methodology follows the scheme displayed in Figure 3.1, where the components indicated in rectangular blocks were fixed, while the components for CNNs and Regressors and Hyperparameters were optimized through grid search. Melt pool images and corresponding KD values were acquired synchronously during laser welding using a coaxial camera and an OCT sensor, respectively. To enable direct comparison, the two data streams were aligned in space and time using linear interpolation based on weld coordinates. Melt pool images were then converted into high-dimensional feature vectors with the aid of a pre-trained CNN. These features were used to train supervised regression models, and model performance was tested

using a three-fold cross-validation scheme.

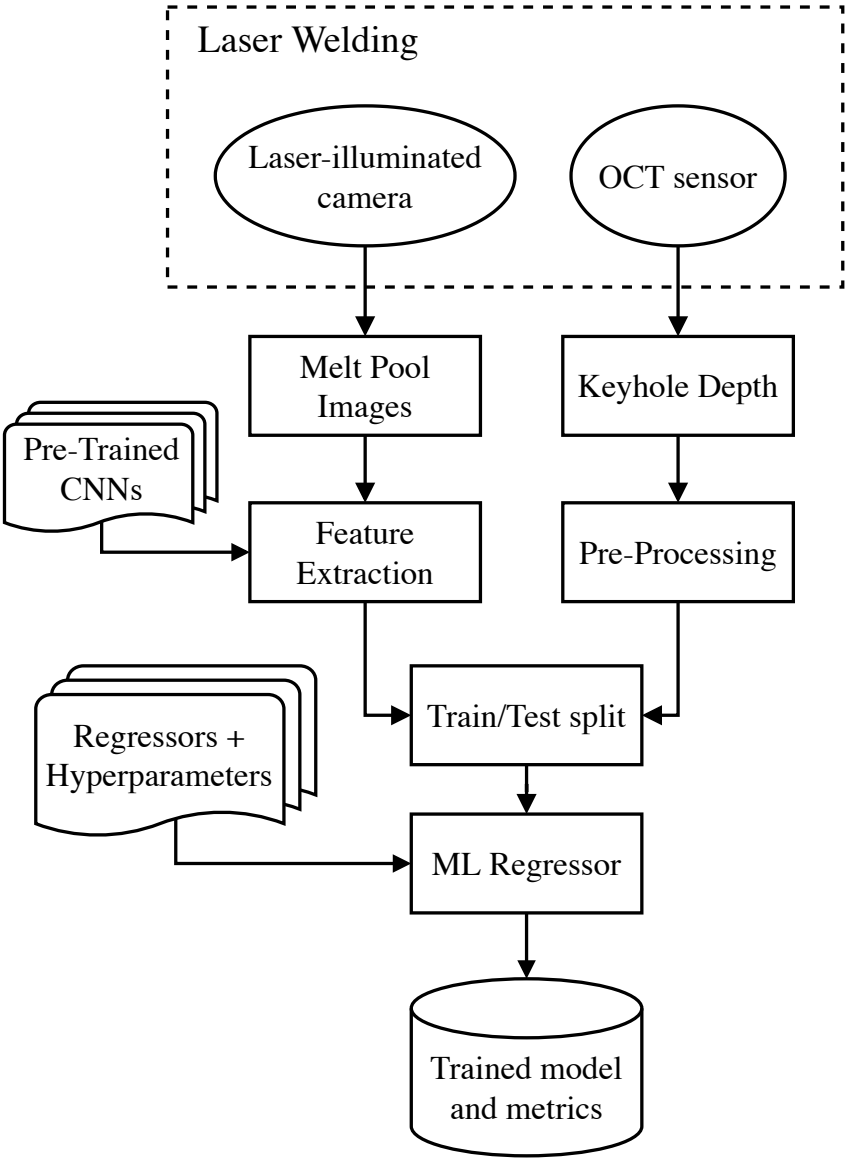

Figure 3.1: Schematic representation of the data processing pipeline for KD prediction.

Experiments involved butt-joint welds with varying laser power to obtain different KD across different runs. In total, 10 welds were performed, with laser power variation according to Table 3.1.

The welding experiments were performed using an IPG YLS-10000MM fiber laser mounted on an ABB robotic arm. Optics were composed with a fiber of 300 µm core

Table 3.1: Different laser power values used in the experiments.

Weld	Laser Power (kW)	Weld (cont.)	Laser Power (kW)
1	4.6	6	5.1
2	4.7	7	5.2
3	4.8	8	5.3
4	4.9	9	5.4
5	5	10	5.5

diameter with 150 mm collimation and 348 mm focusing lenses. In data collection, a coaxially mounted Cavitar C400 welding camera with active laser illumination captured images at 60 Hz, while an IPG LDD-700 integrated with the laser head provided the OCT depth measurements. A high-level schematic of the welding setup is shown in Figure 2.1. The system was augmented with the previously described coaxial camera and OCT sensor for data acquisition.

Table 3.2 summarizes the invariant welding parameters used across ten butt-joint welds. After manually selecting the relevant portions of the dataset (removing non-weld portions), the dataset resulted in 5,608 keyhole images, or approximately 430 images per weld.

Table 3.2: Experimental welding parameters for data collection

Parameter name	Value
Welding speed	$50\,\mathrm{mm\,s^{-1}}$
Weld length	360 mm
Material	ASTM A36 steel
Thickness	5 mm

3.2 Feature Extraction with Pre-trained Models

The data processing methodology employs feature extraction through pretrained CNN models, addressing the inherent data scarcity challenges within LW monitoring domains. The extensive parameter space and experimental setup variability

(characteristic of welding processes) present significant obstacles to constructing comprehensive, general-purpose datasets. Furthermore, limited data accessibility, which stems from proprietary constraints within industrial and research institutions, restricts opportunities for training large robust neural networks from scratch. The data scarcity also unfortunately precludes the creation of a benchmark dataset against which models and approaches could be compared. This challenge provides the foundational motivation for the transfer learning methodology presented in this chapter.

This investigation leverages two established CNN architectures pretrained on large-scale image datasets: DenseNet169 (DN) and EfficientNetB6 (EN). Both architectures demonstrated superior performance in the ImageNet Large Scale Visual Recognition Challenge [85], indicating robust feature learning capabilities and strong generalization potential across diverse visual recognition tasks. Selection criteria prioritized architectures with high representational capacity and computational efficiency.

Following the feature extraction paradigm previously illustrated in Figure 2.7, intermediate convolutional layers are used rather than utilizing final outputs. This strategy repurposes the CNN as a feature encoder, generating high-dimensional representations from learned patterns from the original dataset. Extracting features before the fully connected classification layers preserves hierarchical spatial features and abstract representations inherent to the CNN architecture. These extracted feature vectors then serve as inputs for regression models trained to predict welding process parameters from coaxial imaging data.

3.3 Regression Models and Hyperparameter Optimization

Following the feature extraction step, five regression models have been used to predict the KD. In total, five ML regression models have been chosen: Linear Regressor (LR), GBR, RFR, Elastic Net Regressor (ENR), and SVR. Since the goal of this study is to explore the feasibility of using transfer-learning in melt pool monitoring (broader perspective), rather than improving currently existing methods (narrower perspective), models based on different theories were employed; those different theories were briefly mentioned in section 2.3 such as distance-to-hyperplane (for SVR) and tree ensemble (for RFR and GBR), which is likely to affect the predictive performance. The implementation used in this study is based on the Scikit-learn Python library [86].

Apart from the internal parameter models (e.g. normal vectors in SVR), most ML models also depend on hyperparameters: those values dictate model behavior, size, non-linear combinations, and can affect model performance. Hyperparameter optimization was conducted using grid-search: in this approach, each possible combination of hyperparameters is tested and evaluated, and the parameters and scores for the best tuple of parameters is saved. Grid search has the disadvantage of dramatically increasing model evaluation time, since a vector multiplication is performed: for a hyperparameter space of 3x3x3, 27 models have to be trained and evaluated, and when used in tandem with k-CV, this figure increases by another dimension. Although alternatives to grid-seach exist, for this work the parameter space has been narrowed down in order to balance for development and execution time while also testing combinations that might yield different results. Table 3.3 details the explored parameter spaces for each model. The optimization process considered the number of estimators and maximum depth for GBR and RFR, regu-

larization constants for SVR, and both alpha and L1 ratio parameters for ENR. The LR model was left with the default configuration to serve as a baseline for prediction performance. Finally, the model evaluation, as introduced in section 2.3, has been carried by applying the following scoring functions: R2 (eq. 2.1), MAE (eq. 2.2) and RMSE (eq. 2.4).

Table 3.3: Hyperparameter search space for regression models

Model	**Hyperparameter Space**
GBR	Number of estimators: 100, 250, 350 Maximum depth: 5, 10, default Learning rate: 0.1, 0.01
RFR	Number of estimators: 100, 250, 350 Maximum depth: 5, 10, default Minimum samples to split: 2, 5, 10
SVR	Kernel: RBF Regularization constant: 100, 1000, 1500, 3000, 3500
ENR	Alpha: 1.0, 0.1, 0.01, 0.001 L1 ratio: 0.1, 0.5, 0.9

4 Predicting Laser Power via Acoustic Emission

Laser power represents a critical parameter in LW operations, alongside welding speed [76]. When debris accumulates on optical components including fibre ends and lens windows, the delivered power becomes inconsistent, leading to process discontinuities. Therefore, external monitoring of laser power can enable automated control and alerts when variations in delivered power occur.

A sensing-prediction framework was established based on the premise that LW generates characteristic AE signals. Building upon recent advances in optical microphone technology and progress in ML libraries, this chapter implements and evaluates two parallel approaches for estimating laser power from acoustic measurements.

This chapter presents the methodology and findings for laser power estimation utilizing AE data acquired from an in-process microphone. The organization of this chapter proceeds as follows: the Experimental Setup section details the data acquisition parameters, instrumentation employed, and data conditioning procedures; the Data Processing section outlines the processing pipeline utilized, encompassing dataset preparation, partitioning, and two distinct methods for estimating laser power from AE signatures.

4.1 Experimental Setup

Twenty-two bead-on-plate welding passes were executed on SSAB NV E36 structural steel plates measuring 400 mm x 50 mm x 5 mm, cut using fiber laser technology, as shown in Figure 4.1. Continuous coaxial Argon shielding gas was employed to safeguard both the workpieces and optical system, delivered via the nozzle illustrated in Figure 4.1. Laser power was the sole experimental variable, spanning from 1 kW to 6 kW in 500 W increments. Duplicate runs were conducted at each power setting. This laser power range offers adequate coverage for developing supervised learning models, utilizing acoustic data as input features and laser power as the response variable. The laser power choice was made based on parameters used in the industry for steels of those specifications, as the goal of this project is to reproduce industrial environments and study the feasibility of developing such monitoring system in an industrial context.

An IPG YLS-10000-MM fiber laser was employed as the energy source, mounted on an ABB IRB 4600 6-axis articulated general purpose robot. Optical system inspections were conducted to verify the absence of debris and ensure accurate power delivery. Robot operation was managed through its proprietary scripting interface. Welding speed remained constant at $3\,\mathrm{m\,min^{-1}}$, with sensor and laser activation controlled via synchronized triggering cycles through electrical connections between the computers.

AE monitoring was performed by a Xarion Eta250 Ultra optical microphone, located 140 mm from the process zone, set to acquire data at 2 MHz. Microphone activation started 200 ms before laser activation and stopped 300 ms after the laser deactivation, ensuring valid experimental data. The sensor connected to a QASS Optimizer4D industrial computer featuring a preamplifier for data collection. Signal segments corresponding to stable welding processes were isolated through preprocessing. Via manual examination of time-domain signals, intervals representing

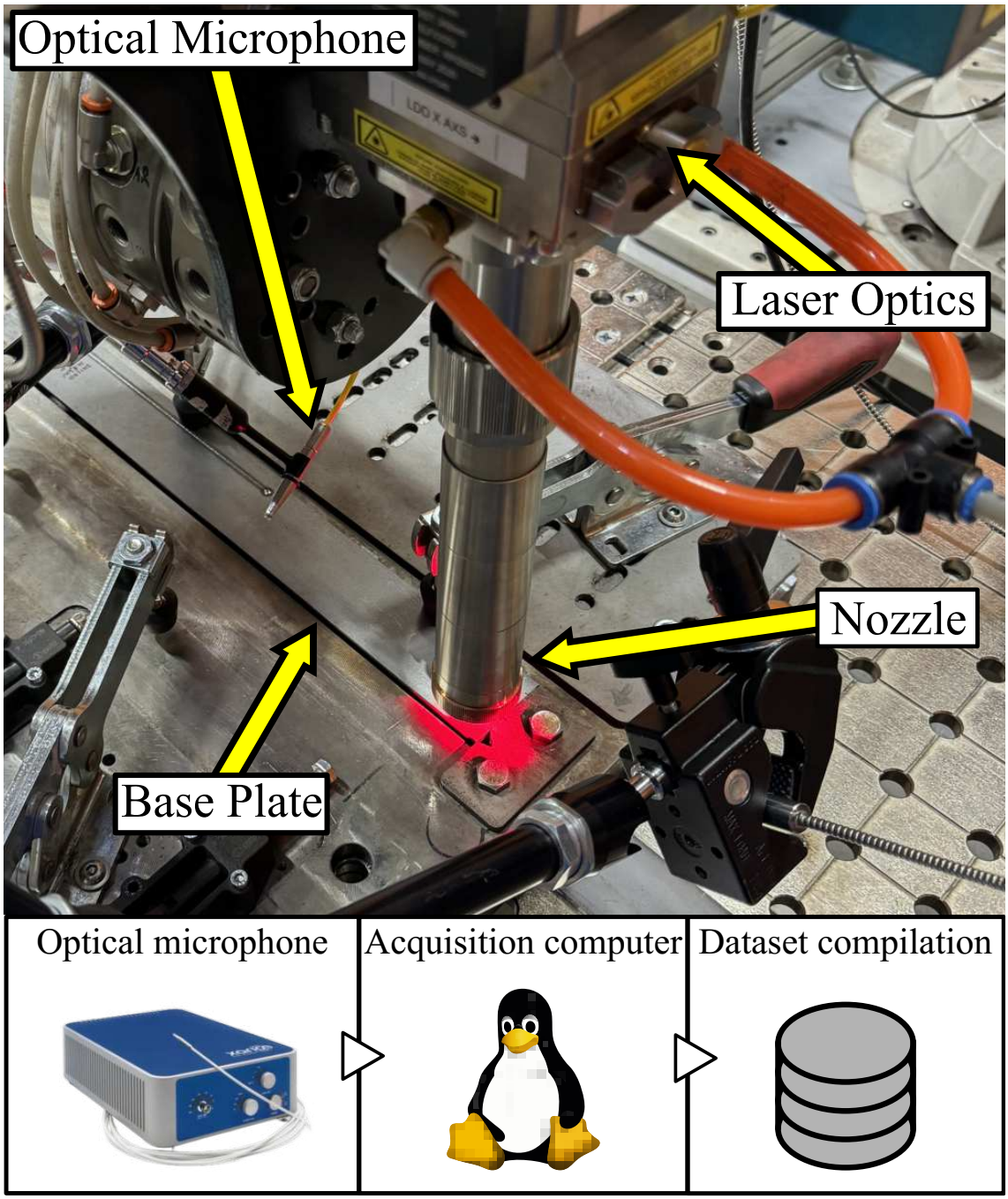

Figure 4.1: Experimental setup for audio processing. Adapted from [23]. Licensed under CC BY 4.0.

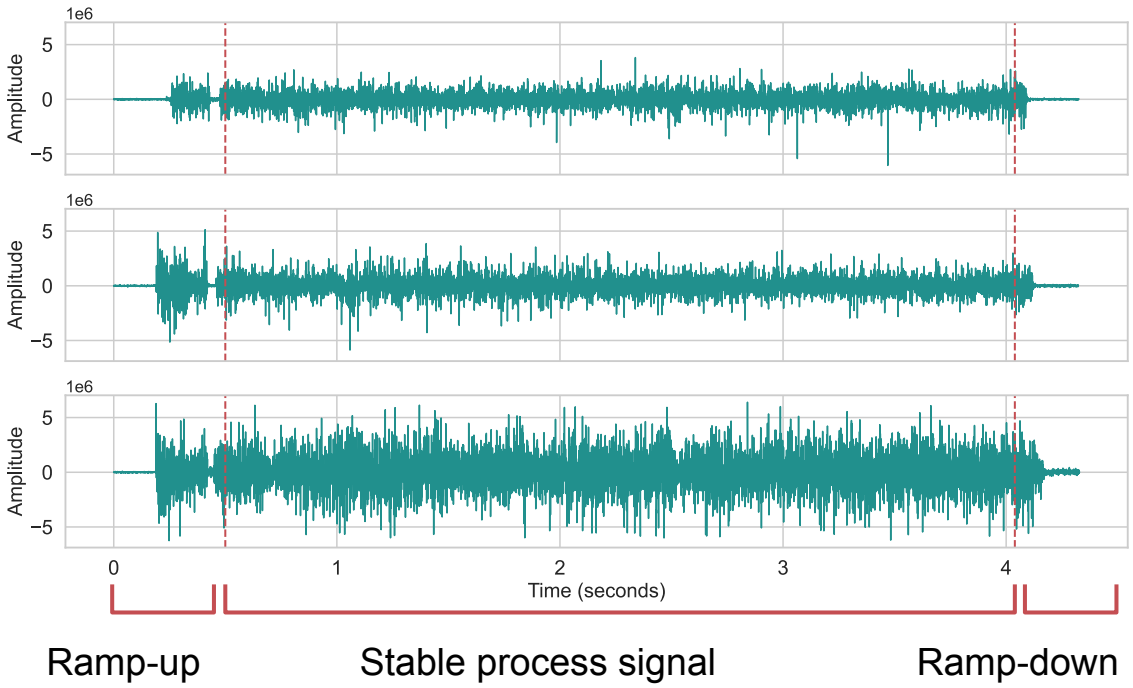

Figure 4.2: Audio selection based signal stability, as it is possible to note from the leading and trailing portions with little or no signal, followed or preceded by a ramp-like signal amplitude change.

steady-state laser power were identified, eliminating the laser's initial ramp-up and final ramp-down segments. This yielded recordings of approximately 3.54 s duration or 7 077 888 samples at 2 MHz sampling rate. Figure 4.2 illustrates the selection boundaries on time domain representations across different recordings.

4.2 Data Processing

Data processing followed the preprocessing stage and includes dataset preparation, model training, and performance assessment. Figure 4.3 illustrates the procedural steps in the data processing methodology.

Beginning with preprocessed signals, a rigid 50:50 train-test partition was implemented (Figure 4.3b), segregating complete experimental runs within respective subsets to prevent data leakage. In this case, the choice of data splitting explicitly

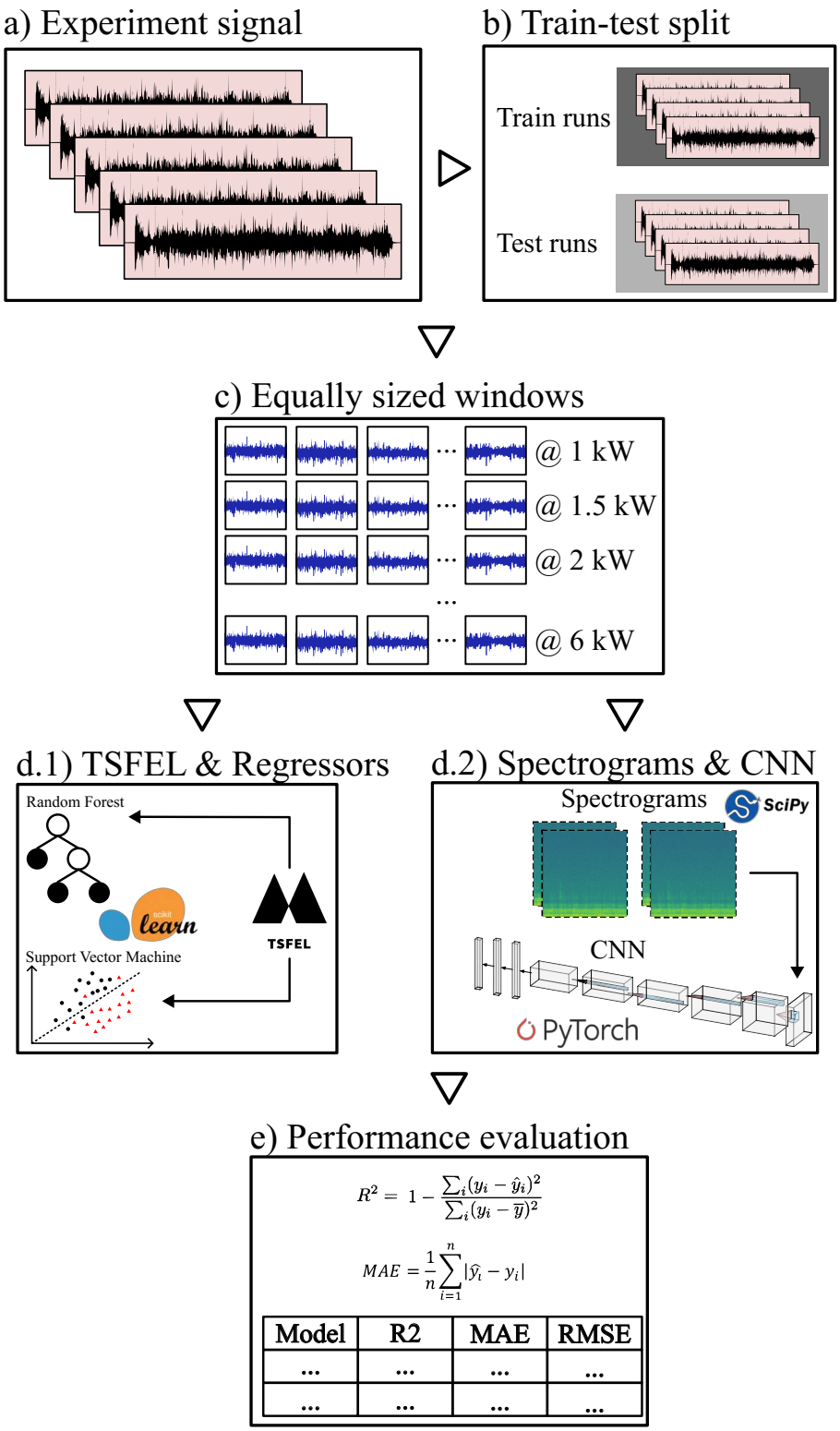

Figure 4.3: Schematic of the proposed data processing pipeline for optical microphone signals. Adapted from [23]. Licensed under CC BY 4.0.

did not follow schemes such as randomized train-test split or k-CV, in order to prevent model biasing. Model bias arises because signals external to the process could be used as a source of information for the prediction model, therefore poisoning the learning process, and this pattern has been observed in similar works presented in the conferences NOLAMP20 and LANE 2024, where I had the opportunity to discuss this approach with other researchers. Poisoning might occur for example, if external noise that is only present in one of the recordings (e.g. 2.5 kW) and the model learns to detect this specific noise to predict 2.5 kW, instead of relying on the process signal itself, as has been proposed in this chapter. Finally, this partitioning produced 11 recordings for training purposes and 11 recordings for testing.

Subsequently, each recording was segmented into 54 equal-sized windows (Figure 4.3c), representing brief observations of the LW process at approximately 66 ms duration. Windowing served to generate multiple observations from individual experimental runs, transforming single bead-on-plate acquisitions into numerous data points. Additionally, this strategy better approximates real-life scenarios, where monitoring algorithms would process discrete signal segments. Two window-level processing strategies were then developed: the first leverages the Time Series Feature Extraction Library (TSFEL) (Figure 4.3d.1), while the second converts signals into spectrograms and employs a CNN to estimate laser power from spectrogram representations (Figure 4.3d.2). Finally, both methods were evaluated using standard regression metrics (Figure 4.3e).

4.2.1 TSFEL-based Prediction

The approach shown in Figure 4.3d.1 employs TSFEL as its initial step. Built as an open-source Python library, TSFEL calculates a series of metrics from input time series [87]. According to the official documentation, the library encompasses four principal feature categories: temporal features capture time-dependent pat-

terns including trends, cycles, and temporal correlations while remaining sensitive to observation ordering; statistical features characterize data distributions through descriptive metrics such as mean, variance, skewness, and kurtosis, independent of temporal ordering; spectral features analyze frequency-domain content via Fourier and wavelet transformations to reveal periodicities, harmonics, and cyclical components; and fractal features quantify complexity and self-similarity across multiple scales through measures such as fractal dimension, capturing intricate patterns characteristic of complex systems.

The short windows (Figure 4.3c) were processed by first extracting all available metrics in TSFEL, followed by using the built-in utility for eliminating redundant and low-variance features. In this procedure, an initial feature set of 312 metrics was computed for each window, and subsequently reduced to 105 features through the library's integrated redundancy elimination procedure.

The described feature extraction step was followed by the training and assessment of 2 ML regression models available within the Scikit-learn Python library [86]: SVR and RFR. These were selected based on widespread adoption and implementation simplicity. The dataset containing extracted features was guaranteed to be standardized (z-score standardization) prior to training and evaluation.

4.2.2 CNN-based Prediction

To offer an alternative to feature engineering and traditional ML approaches, a deep learning strategy was developed, illustrated in Figure 4.3d.2. Within this implementation, each window undergoes transformation into a spectrogram via the Short Time Fourier Transform (STFT) utility from the SciPy Python library [88]. Configuration parameters included a "Hanning" window of 4096-point width with 75 % overlap between successive windows. Spectrograms were subsequently converted to dB scale and cropped from their original 2049x131 dimensions to 2048x128 pixel

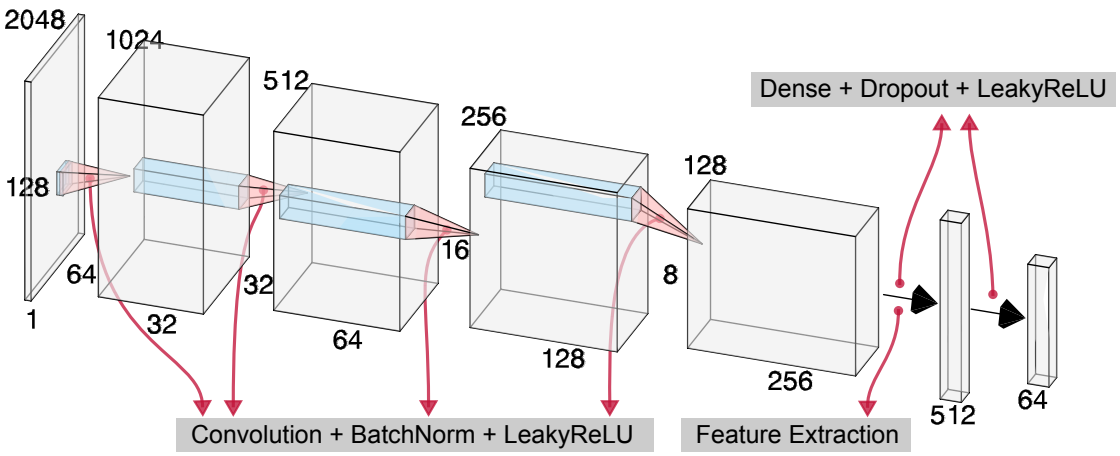

Figure 4.4: Abstract CNN architecture exposing the data dimensionality between operational blocks.

2-D arrays for streamlined downstream processing.

Two-dimensional arrays share the identical data structure as images, rendering image processing algorithms a feasible option. As CNNs have demonstrated great performance in computer vision tasks, a CNN was implemented using the PyTorch framework [89] and its architecture is depicted in Figure 4.4. In a higher level of detail, the Python code in Listing 1 is the PyTorch definition, where it is possible to note the data compression blocks, composed of `Conv2d`, `BatchNorm2d` and `LeakyReLU`; and the regression head, composed of `Linear`, `Dropout` and `LeakyReLU`. Apart from the prediction data flow, a feature extraction function was created by applying the encoder step on the input image and flattening the encoder output to 1-D. The flattened vector is a feature map and was used as a comparison for features extracted by TSFEL. Training utilized the Adam optimizer with a 0.001 learning rate across 100 epochs. MSE serves as the loss function, appropriate for continuous-valued variable prediction.

Table 4.1: Parameters for STFT applied to the signal windows.

Window Width	4096
Window Overlap	75 %
Spectrogram Size	2049x131 to 2048x128
Spectrogram Scale	dB

Table 4.2: Parameters for CNN training.

Loss Function	MSE
Optimizer	Adam
Learning Rate	0.001
Epochs	100

4.2.3 Model Evaluation

The model was trained using the training subset and evaluated using the test subset, depicted in Figure 4.3b. With the intent of providing a holistic perspective, low-dimensionality scatter plots were created for a qualitative analysis of the feature representations of both the TSEFL and the CNN-based method.

Given the data split, after training and running the inferences, the predictions have been compared using the scoring functions R^2, MAE and RMSE, described in Equations 2.1, 2.2, and 2.4, respectively. The metrics for the different approaches were compiled into a table, which has been complemented by a prediction plot for the best predictive model, together with error residuals across the dataset. Since this work is among the initial in AE-based laser power prediction for LW, there is no well defined baseline for performance. Therefore, the results can offer insights into what is possible with this setup, and also offer ideas for what can be further developed.

Listing 1 PyTorch implementation for the spectrogram-based CNN model.

```
class SpectrogramCNN(nn.Module):
    def __init__(self, input_channels=1):
        super().__init__()

        # Convolutional encoder
        self.encoder = nn.Sequential(
            # 2048×128 → 1024×64
            nn.Conv2d(input_channels, 32,
                kernel_size=4, stride=2, padding=1),
            nn.BatchNorm2d(32),
            nn.LeakyReLU(0.2),

            # 1024×64 → 512×32
            nn.Conv2d(32, 64, kernel_size=4, stride=2, padding=1),
            nn.BatchNorm2d(64),
            nn.LeakyReLU(0.2),

            # 512×32 → 256×16
            nn.Conv2d(64, 128, kernel_size=4, stride=2, padding=1),
            nn.BatchNorm2d(128),
            nn.LeakyReLU(0.2),

            # 256×16 → 128×8
            nn.Conv2d(128, 256, kernel_size=4, stride=2, padding=1),
            nn.BatchNorm2d(256),
            nn.LeakyReLU(0.2)
        )

        self.flatten_dim = 256 * 128 * 8

        # Regression head
        self.regressor = nn.Sequential(
            nn.Flatten(),
            nn.Linear(self.flatten_dim, 512),
            nn.LeakyReLU(0.2),
            nn.Dropout(0.3),
            nn.Linear(512, 64),
            nn.LeakyReLU(0.2),
            nn.Linear(64, 1)
        )

    def forward(self, x):
        x = x.unsqueeze(1)
        features = self.encoder(x)
        output = self.regressor(features)
        return output.squeeze(-1)
```

5 Results and Discussion

The goal of the current chapter is to present and analyze the results of the experiments and data processing methodology proposed in chapters 3 and 4. Model performance has been analyzed and compared against the ground-truth data collected during the experimentation phase. Furthermore, the performance comparison across the datasets have been displayed visually and will be discussed in the respective sections. This chapter also aims at discussing limitations of the proposed methodology, which serves as foundation to the conclusions drawn in chapter 6 and the future work proposals in chapter 7.

5.1 Coaxial Images

The experimental proposal and methodology for chapter 3 were executed: LW was performed at a range of laser powers, with a simultaneous monitoring from a coaxial laser-illuminated camera providing a top-view of the melt pool and OCT sensor configured to track the KD (see Figure 3.1), the data has been organized according to good practices and the images were saved in the TIF format while the time series related to the KD measurement was saved in CSV. In Figure 5.1, it is possible to observe differences in the coaxial melt pool image in a single experiment. This variation is caused partly by the liquid metal-vapor interaction, and by other thermodynamical phenomena that are combined into a complex behavior.

Prediction models consisting of pre-trained CNNs as feature extractors in tandem

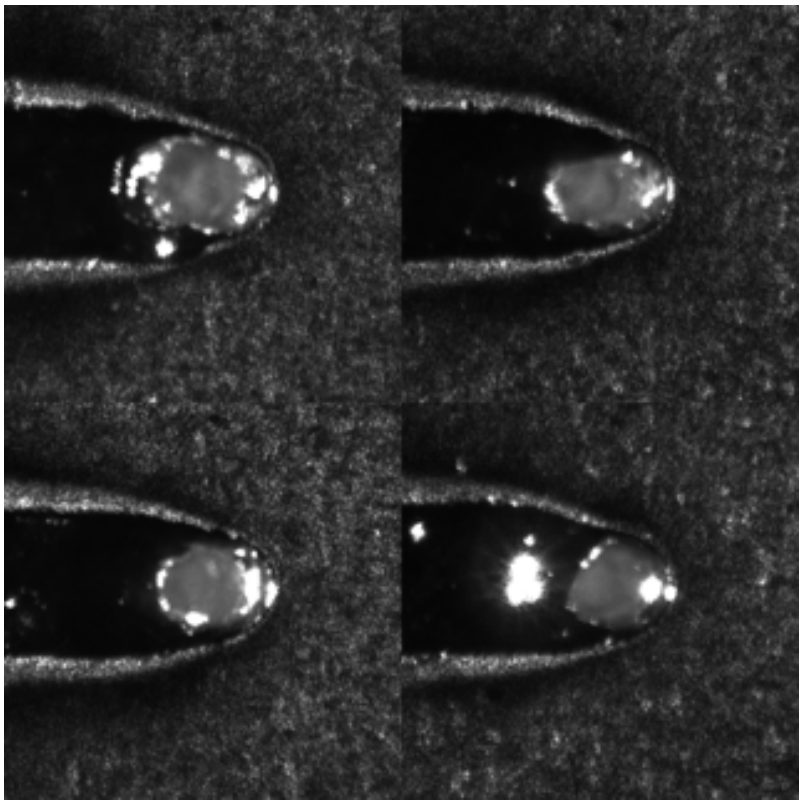

Figure 5.1: Four melt pool images collected during a single LW experiment, at different timepoints.

with ML regressors were evaluated and summarized in Table 5.1, which presents the 5-fold k-CV performance metrics for the proposed combinations of feature extractors and regression models.

Table 5.1: Performance metrics for all model combinations (best and second-best scores in bold)

Model Code	**MAE** (µm)	**MAPE** (%)	$\mathbf{R^2}$
EN_LR	135.21	3.00	0.56
EN_GBR	148.86	3.30	0.47
EN_RFR	162.60	3.61	0.38
EN_ENR	134.44	2.98	0.57
EN_SVR	138.77	3.08	0.54
DN_LR	131.93	2.92	0.58
DN_GBR	148.31	3.28	0.48
DN_RFR	163.82	3.63	0.37
DN_ENR	**130.78**	**2.89**	**0.59**
DN_SVR	**131.40**	**2.90**	**0.59**

It is possible to observe from the table that the largest difference in performance was due to the downstream regressor, while the difference in CNNs models was smaller (e.g. table lines EN_LR and DN_LR, or EN_ENR and DN_ENR). In a quantitative analysis, the approach of DenseNet169 as feature extractor with ENR as regressor achieved the highest performance with an R^2 of 0.59, MAE of 130.78 µm and MAPE of 2.89 %. Demonstrating similar predictive power, the DenseNet169 combined with SVR resulted in an R^2 of 0.59, MAE of 131.40 µm and MAPE of 2.90 %.

Offering a graphical analysis of the performance across the entire range of data, Figure 5.2 portrays the performance of the best aforementioned models on unseen data, by using the randomized k-CV folds as a prediction, and feeding thus the entire dataset through the series of models. In this Figure one can observe that the majority of data points (87 % of observations) fall within the 5 % error bounds, with no systematic bias observed across different depth ranges. The achieved 2.89 % MAPE represents high relative accuracy for most industrial applications.

Notwithstanding, the obtained R^2 value (ideally equal to 1) offers a different perspective and indicates that the explained variance of the analyzed models was found wanting and that, in this context, additional explanatory variables or more sophisticated models could further improve performance. This issue may also stem from experimental setup limitations (which has also been discussed in 2.2) or the limited accuracy of KD monitoring via coaxial imaging. Further study may involve experimenting with a larger dataset of melt pool images and OCT measurements and training more powerful models (especially domain-specific models), while validating further performance improvements.

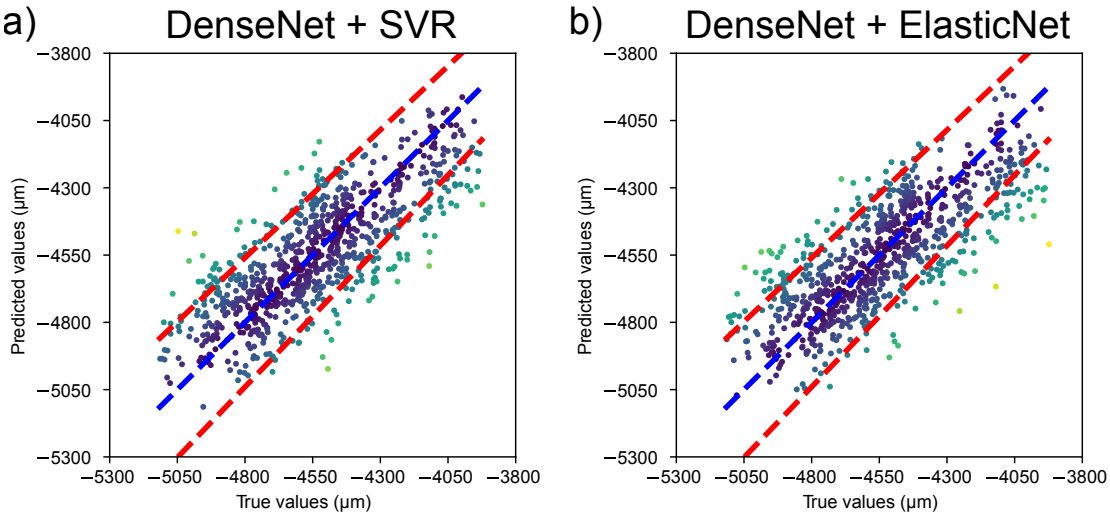

Figure 5.2: Comparison between real and predicted KD. The real regression is represented by a blue dashed line, and the 5 % error boundary is represented by a red dashed line.

Finally, another potential issue from this data collection and validation approach is that the dataset was divided via randomized train-test split, but the observations were not independent: in this methodology, one detected problem was that the images and OCT readings from a single weld were used both for training and for testing, which allows a biased model to be trained, as explained and mitigated in Chapter 4. The stricter separation of data in the aforementioned chapter reduces the bias resulting from the model learning on confounding factors (e.g. base plate

color, difference in external illumination from the camera) instead of learning specific process information (e.g. melt pool and keyhole geometry). This information external to the process allows the model to deviate from the intended optimization, which is model capable of generalizing across coaxial images in LW.

An execution time analysis was additionally carried out, and the inference time has been measured, including both the time for feature extraction and for regression. Averaging the entire processing time, the approach of DenseNet169 as feature extractor with ENR yielded circa 30 ms of processing time per image, which can be inverted and translated into approximately 30 frames per second of image inference, which is considered as sufficient in feedback control systems. It is worth pointing out that the pipeline has been implemented in an equivalent of a widely available and cost-accessible computer system (as of 2025), without relying on HPC systems or high-end discrete graphics cards, suggesting that the proposed approach could be transferred into industrial computers. However, it is also worth mentioning that there are several optimization directions for implementations in constrained computer hardware (such as embedded systems and low-power devices) that could in a near future be deployed into an integrated camera solution. Such optimizations can be carried out in different levels: in a lower Level of Abstraction (LoA), performance improvements might be witnessed by implementing the methodology into lower-level programming languages (e.g. C++ or Rust); in a middle LoA, weight-pruning can be done to reduce redundant model components; in a higher LoA, a more optimized CNN can be trained, in end-to-end that directly outputs the value of interest.

5.2 Acoustic Emission

The dataset was collected, and Figure 5.3 displays a top view showing the welded beads alongside the acoustic emission signal recorded during the first produced bead. This figure demonstrates that bead width and audio signal amplitude both increase

as laser power rises. Since laser power was the sole parameter varied between experiments, the acoustic signals originating from the melt pool and keyhole dynamics exhibit a clear correlation with the applied power. Following the windowing procedure in the data processing pipeline, 1188 audio segments were obtained, with each segment comprising 131 072 samples. These segments were divided equally into training and testing subsets of 594 windows each. Feature extraction, model training, and evaluation were then conducted on each windowed segment as outlined in section 4.2.

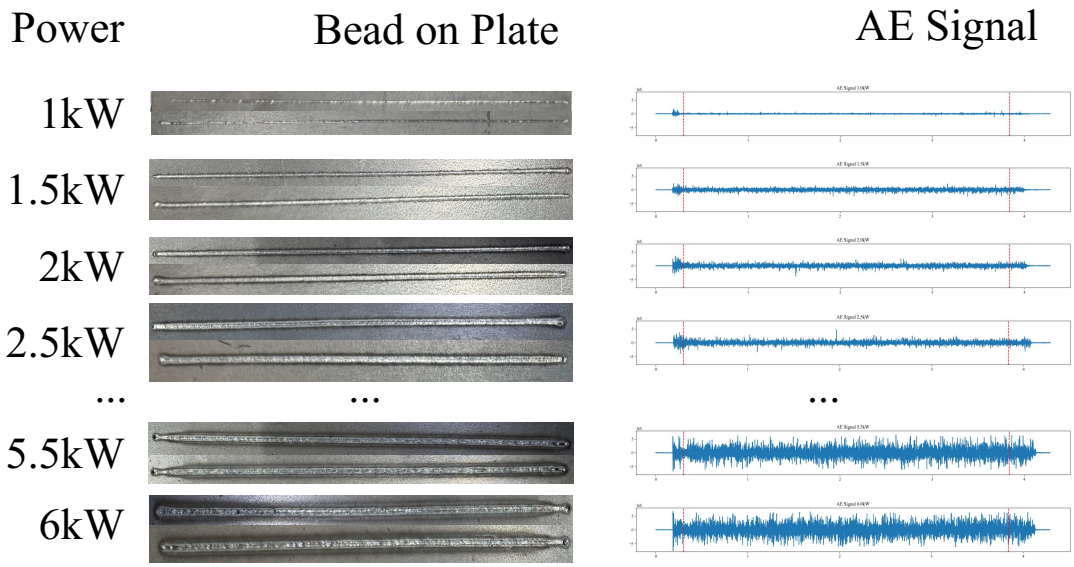

Figure 5.3: Bead on plates and time-domain audio

To provide a qualitative assessment of how effectively the learned features epitomize data patterns, visualizations in reduced-dimensional space were created through Principal Component Analysis (PCA). Figure 5.4 presents PCA visualizations for features derived through TSFEL and through the CNN architecture. Within these visualizations, point colors represent the laser power setting during signal capture. Through a visual observation, data points associated with comparable laser power levels exhibit spatial clustering in the 2D PCA representation for both feature sets, indicating that each feature extraction methodology has successfully encoded vari-

ance attributable to laser power variations.

The variance explained by the first two principal components (PC1 and PC2) in the TSFEL feature space reached approximately 25 %. Analysis revealed that PC1 was predominantly composed of "absolute energy" and "mean spectrogram coefficient at 580 645 Hz" [1], implying increased sensitivity around the 580 kHz spectral region to laser power changes under these experimental conditions. For PC2, the primary contributors were wavelet entropy, median frequency, and neighborhood peaks features. In contrast, the CNN-derived feature space showed PC1 and PC2 accounting for approximately 73 % of the total variance.

Table 5.2: Model metrics for the different methodologies.

Method	**R^2**	**MAE** (kW)	**RMSE** (kW)
TSFEL + RFR	0.918	0.246	0.452
TSFEL + SVR	0.926	0.309	0.429
CNN	0.913	0.339	0.458

Performance metrics including R^2, MAE, and RMSE for each modeling approach are compiled in Table 5.2. The method using TSFEL combined with SVR (indicated as TSFEL+SVR in the table) delivered superior performance, achieving an R^2 of 0.926, MAE of 0.309 kW, and RMSE of 0.429 kW. The approach utilizing TSFEL combined with RFR (indicated as TSFEL+RFR in the table) exhibited comparable although marginally lower performance metrics, reaching an R^2 of 0.918. The CNN-based method produced an R^2 of 0.913. Given that all R^2 values exceed 0.9, these findings confirm that each methodology successfully captures a significant proportion of laser power variance present in the acoustic emission data.

Figure 5.5 visually displays a comparison between the ground-truth experimental data and the values predicted by the best-performing regression model in Table 5.2, TSFEL combined with SVR. In the upper section of the figure, the prediction for

[1] Feature definitions from [87]

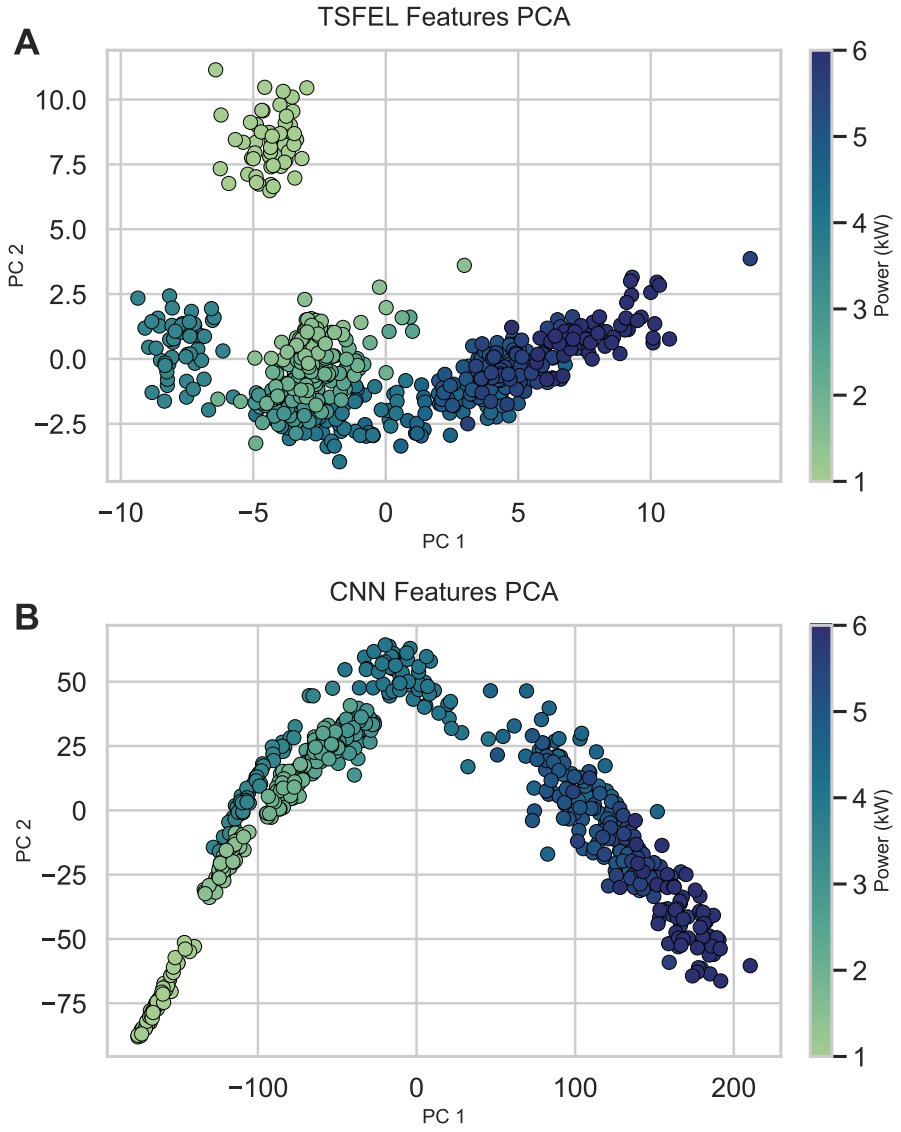

Figure 5.4: Plots in PCA space using features extracted by (a) TSFEL and (b) custom CNN model, The observation points have been colored according to the laser power.

each observation window is a dot and the juxtaposed dashed line indicates an ideal prediction; it is possible to observe that the predictions follow the trend of true values, with few outliers with larger error magnitude. Concomitantly, the lower section of the plot displays box plots of the prediction residuals sharing the same horizontal axis for ground truth data. This allows for a clearer analysis of the distribution of the error, otherwise difficult to observe in a scatter plot where points occlude others in the same region. Finally, the distribution of predictions shows close alignment with this ideal line, suggesting a relatively consistent error pattern throughout the complete dataset.

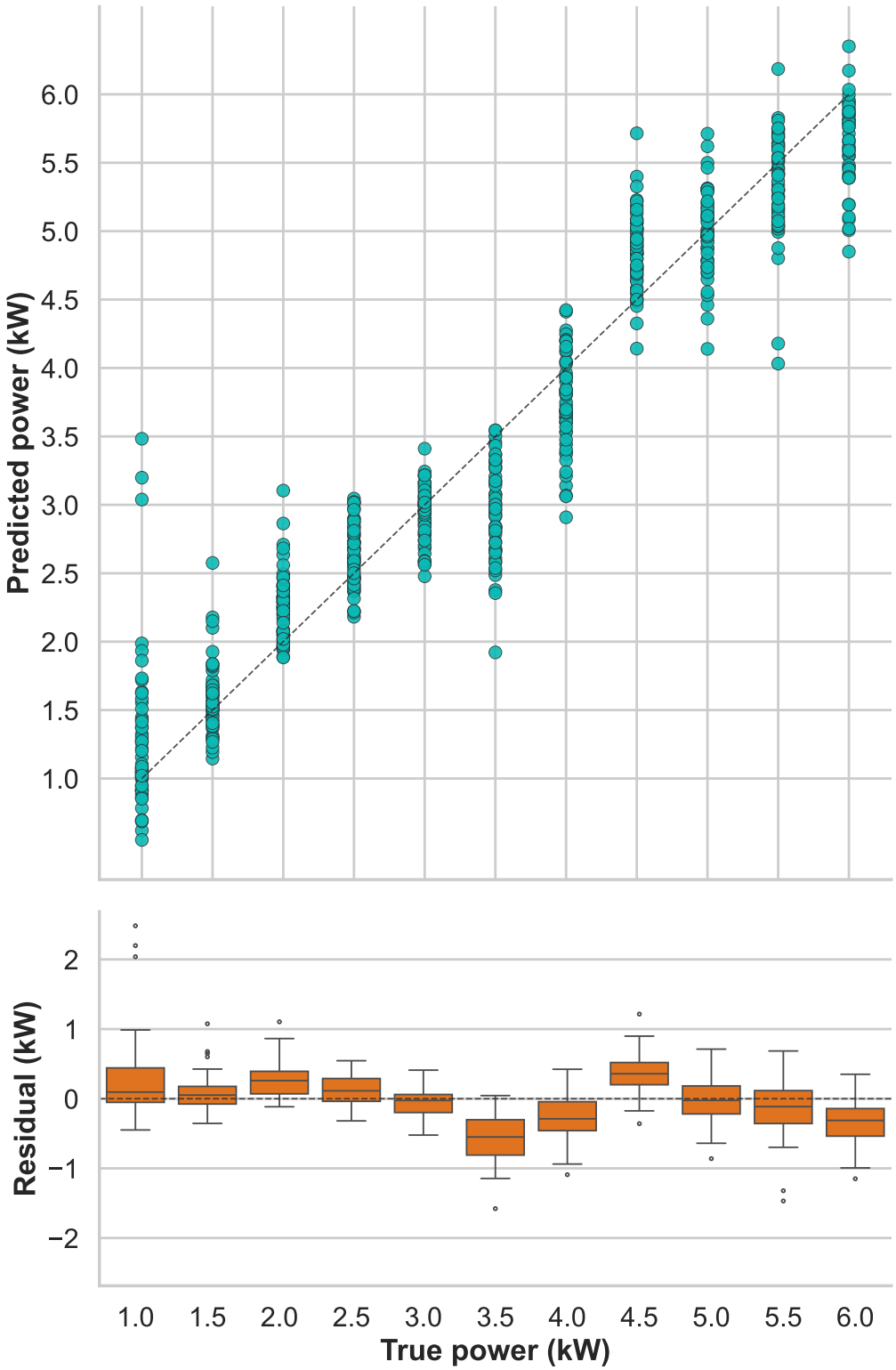

Figure 5.5: Comparison between ground-truth data and predictions of the model consisting of TSFEL and SVR.

6 Conclusions

This work developed two in-process monitoring methodologies for laser welding processes. The first methodology used coaxial imaging for keyhole depth prediction, while the second used acoustic emission signals for laser power estimation. Both approaches address practical constraints common in industrial welding environments, including limited training data availability and computational resource constraints.

A camera-based monitoring system derived from transfer learning from pretrained CNN architectures for feature extraction is proposed and validated, as presented in Chapter 3. The methodology combined DenseNet169-based feature extraction with SVR, achieving a MAE of 130.78 µm (2.89% MAPE) with a real-time-compatible prediction time on general purpose hardware. These results indicate that coaxial cameras can be applied for an keyhole depth estimation in the absence of direct OCT measurements in the context of structural steel LW. The demonstrated performance suggests potential for deployment in complementary monitoring tasks including visual inspection and seam tracking.

However, limitations were observed. The R^2 value of 0.59 indicates that substantial portion of data variance was not explained by the proposed model. This finding suggests either that alternative data processing strategies (e.g. end-to-end CNN training with larger datasets) may yield improved performance, or more fundamentally, that a single coaxial camera alone provides insufficient information for high-accuracy depth estimation under the studied setup. This limitation aligns with

observations from the literature review (Chapter 2), in which the comparison between 2D and 3D measuring system was made. Additionally, the train-test splitting methodology employed may have introduced model bias, as the observation points were randomized and information from the same experimental run was available both in training and testing. Future work should adopt strategies such as the ones used in Chapter 4 to provide more robust performance estimates.

Chapter 4 explored a complementary monitoring approach utilizing high-frequency AE data acquired through an optical microphone. The investigation evaluated three processing pipelines across two distinct methodologies. The optimal configuration, which combined TSFEL feature extraction with SVR, achieved an R^2 of 0.926 and MAE of 0.309 kW, demonstrating robust correlation between acoustic signatures and laser power variations. These results validate AE monitoring as an effective modality for process surveillance, enabling detection of power delivery instabilities that may arise from optical contamination or different factors.

The successful implementation of both image-based and acoustic-based monitoring systems demonstrates that cost-effective sensor configurations, when paired with appropriate ML architectures, can achieve monitoring performance suitable for industrial deployment. These approaches offer particular value in cost-sensitive applications where more expensive sensing technologies such as OCT cannot be applied. The complementary characteristics of visual and acoustic data streams suggest significant potential for multi-sensor fusion architectures. Such fusion models could leverage both (i) the geometric aspects of the melt pool and top view of the keyhole, and (ii) high frequency signal emitted by the fusion of the material during LW. This technique can also be transferred to other laser-based processes such as Additive Manufacturing, which are idiosyncratic in different directions (e.g. material delivery, layer height). As a further reflection, establishing standardized performance benchmarks and acceptable tolerance thresholds within industrial con-

texts would facilitate meaningful comparisons across monitoring methodologies and accelerate technology transfer to production environments.

In a broader context, the past century has seen multiple enabling technologies (such as the development of lasers, computers, digital cameras and communication protocols). The twenty-first century presents a natural evolution towards refining existing technologies for increased efficiency and reduced error margins, while simultaneously developing advanced data processing and automation systems capable of addressing increasingly complex challenges. These challenges encompass demographic transformations, environmental sustainability demands, and the imperative for resource-efficient manufacturing processes.

Within this context, the monitoring approaches presented in this thesis contribute to the ongoing advancement of intelligent manufacturing systems. The integration of machine learning with cost-effective sensing modalities demonstrates a pathway toward democratizing advanced process control capabilities across diverse industrial scales and economic contexts.

7 Future Work

Several future directions could extend the research presented in this thesis. The generalization of the proposed methodologies could be validated by extending their application to different materials, joint geometries, and welding setups. The creation of larger and more diverse datasets would enable the development and training of more advanced deep learning architectures, potentially improving predictive accuracy and robustness.

This work has developed methodologies that use isolated information (camera, microphones) to predict specific parameters. From a sensor fusion perspective, it is sensible to develop algorithms that combine multi-modal analysis into a model capable of delivering higher predictive performance and also be robust against process signal discontinuities by using a different data source unaffected by a specific noise.

The next logical step is to integrate these predictive models into a closed-loop control system. Such a system could use the real-time keyhole depth and laser power predictions to dynamically adjust process parameters, thereby preventing defects and ensuring consistent weld quality. Finally, real-world validation studies in an industrial production environment are essential to confirm the long-term stability, reliability, and economic benefit of these monitoring systems.

References

[1] B. Zhou and Q. Wu, "Decomposition-based bi-objective optimization for sustainable robotic assembly line balancing problems", *Journal of Manufacturing Systems*, vol. 55, pp. 30–43, Apr. 1, 2020, ISSN: 0278-6125. DOI: 10.1016/j.jmsy.2020.02.005.

[2] P. J. Gisi, *The Dark Factory and the Future of Manufacturing: A Guide to Operational Efficiency and Competitiveness*. Taylor and Francis, 2024, 252 pp. DOI: 10.4324/9781032688152.

[3] S. Krüger, A.-K. Lange, and C. Jahn, "Development of Global Seabound Mobility", in *Green Energy and Technology*, vol. Part F3771, Springer Science and Business Media Deutschland GmbH, 2025, pp. 89–122. DOI: 10.1007/978-3-031-62411-7_5.

[4] T. Falcao, "Taxing Carbon Emissions from International Shipping", *Intertax*, vol. 47, no. 10, Oct. 1, 2019, ISSN: 0165-2826.

[5] M. dominic and C. Nandakumar, "Environmental impact of Non Metallic Hull Ships", in *2012 International Conference on Green Technologies (ICGT)*, Dec. 2012, pp. 307–312. DOI: 10.1109/ICGT.2012.6477991.

[6] T. Lee and H. Nam, "A Study on Green Shipping in Major Countries: In the View of Shipyards, Shipping Companies, Ports, and Policies", *The Asian Journal of Shipping and Logistics*, vol. 33, no. 4, pp. 253–262, Dec. 1, 2017, ISSN: 2092-5212. DOI: 10.1016/j.ajsl.2017.12.009.

[7] H. Shim, Y. H. Kim, J.-P. Hong, D. Hwang, and H. J. Kang, "Marine Demonstration of Alternative Fuels on the Basis of Propulsion Load Sharing for Sustainable Ship Design", *Journal of Marine Science and Engineering*, vol. 11, no. 3, 2023, ISSN: 2077-1312. DOI: 10.3390/jmse11030567.

[8] M. Ivanov et al., "Experimental and numerical investigation on cold cracking susceptibility of naval grade high strength steel welds for lightweight shipbuilding structures", *International Journal of Lightweight Materials and Manufacture*, vol. 7, no. 4, pp. 597–613, Jul. 1, 2024, ISSN: 2588-8404. DOI: 10.1016/j.ijlmm.2024.03.004.

[9] C. Churiaque, J. M. Sánchez-Amaya, Ö. Üstündağ, M. Porrua-Lara, A. Gumenyuk, and M. Rethmeier, "Improvements of hybrid laser arc welding for shipbuilding T-joints with 2F position of 8 mm thick steel", *Optics & Laser Technology*, vol. 143, p. 107 284, Nov. 1, 2021, ISSN: 0030-3992. DOI: 10.1016/j.optlastec.2021.107284.

[10] C. Pyo, J. Kim, and H. Oh, "Multi-layered heat source model for fiber laser welding of cryogenic steel", *Journal of Ocean Engineering and Science*, Aug. 8, 2025, ISSN: 2468-0133. DOI: 10.1016/j.joes.2025.08.003.

[11] S. Katayama, "12 - Defect formation mechanisms and preventive procedures in laser welding", in *Handbook of Laser Welding Technologies*, ser. Woodhead Publishing Series in Electronic and Optical Materials, S. Katayama, Ed., Woodhead Publishing, Jan. 1, 2013, pp. 332–373. DOI: 10.1533/9780857098771.2.332.

[12] N. Kumar, V. V. Pamarthi, J. Moffat, and I. Masters, "In-depth evaluation of laser welding of thick busbar to 21700 Li-ion cell terminal for electric supercar vehicle battery pack", *Journal of Materials Research and Technology*, vol. 33,

pp. 3058–3067, Nov. 1, 2024, ISSN: 2238-7854. DOI: 10.1016/j.jmrt.2024.10.012.

[13] Q. Song et al., "Design and mechanism research of online control and monitoring system for laser welding quality", *Materials & Design*, vol. 259, p. 114 849, Nov. 1, 2025, ISSN: 0264-1275. DOI: 10.1016/j.matdes.2025.114849.

[14] J. Stavridis, A. Papacharalampopoulos, and P. Stavropoulos, "Quality assessment in laser welding: A critical review", *The International Journal of Advanced Manufacturing Technology*, vol. 94, no. 5–8, pp. 1825–1847, Feb. 2018, ISSN: 0268-3768, 1433-3015. DOI: 10.1007/s00170-017-0461-4.

[15] L. Alzubaidi et al., "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions", *Journal of Big Data*, vol. 8, no. 1, p. 53, Mar. 31, 2021, ISSN: 2196-1115. DOI: 10.1186/s40537-021-00444-8.

[16] G. Masinelli, "Toward Autonomous Laser Manufacturing: Solving Key Challenges in Optimization, Monitoring, and Control", EPFL, 2025. DOI: 10.5075/epfl-thesis-10643.

[17] K. Nam and H. Ki, "One camera-based laser keyhole welding monitoring system using deep learning", *Journal of Manufacturing Processes*, vol. 104, pp. 17–27, Oct. 27, 2023, ISSN: 1526-6125. DOI: 10.1016/j.jmapro.2023.08.056.

[18] M. Klaiber, M. Hartmann, J.-P. Hermani, A. Jahn, and A. Michalowski, "Multimodal sensor fusion with SWIR imaging and audio for inline gas-tightness monitoring in laser-welded bipolar plates", *Optics & Laser Technology*, vol. 192, p. 113 441, Dec. 1, 2025, ISSN: 0030-3992. DOI: 10.1016/j.optlastec.2025.113441.

[19] X. Xiao, X. Liu, M. Cheng, and L. Song, "Towards monitoring laser welding process via a coaxial pyrometer", *Journal of Materials Processing Technol-*

ogy, vol. 277, p. 116 409, Mar. 1, 2020, ISSN: 0924-0136. DOI: 10.1016/j.jmatprotec.2019.116409.

[20] D. Wu, P. Zhang, H. Shi, Q. Lu, Z. Yu, and Y. Huang, "Advancements and prospects of OCT-enabled all-process monitoring and inline quality assurance in laser keyhole welding: A critical review", *Journal of Manufacturing Processes*, vol. 152, pp. 1179–1203, Oct. 30, 2025, ISSN: 1526-6125. DOI: 10.1016/j.jmapro.2025.08.063.

[21] I. S. genannt Wersborg, T. Bautze, F. Born, and K. Diepold, "A cognitive approach for a robotic welding system that can learn how to weld from acoustic data", in *2009 IEEE International Symposium on Computational Intelligence in Robotics and Automation - (CIRA)*, Dec. 2009, pp. 108–113. DOI: 10.1109/CIRA.2009.5423224.

[22] H. H. L. Núñez, L.-W. Hsu, K. S. B. Ribeiro, A. Salminen, and W. M. Bessa, "In-situ monitoring and online prediction of keyhole depth in laser welding by coaxial imaging", *Procedia CIRP*, 13th CIRP Conference on Photonic Technologies [LANE 2024], 15-19 September 2024, Fürth, Germany, vol. 124, pp. 793–796, Jan. 1, 2024, ISSN: 2212-8271. DOI: 10.1016/j.procir.2024.08.227.

[23] H. H. Libutti-Núñez, L.-W. Hsu, S. Parchegani, K. S. B. Ribeiro, W. Moreira Bessa, and A. Salminen, "Evaluation of Acoustic Emission as a Predictor of Laser Power in Laser Welding", *IOP Conference Series: Materials Science and Engineering*, vol. 1332, no. 1, p. 012 041, Aug. 2025, ISSN: 1757-899X. DOI: 10.1088/1757-899X/1332/1/012041.

[24] T. H. Maiman, "Stimulated Optical Radiation in Ruby", *Nature*, vol. 187, no. 4736, pp. 493–494, Aug. 1960, ISSN: 1476-4687. DOI: 10.1038/187493a0.

[25] "Sixty years of lasers", *Nature Reviews Physics*, vol. 2, no. 5, pp. 221–221, May 2020, ISSN: 2522-5820. DOI: 10.1038/s42254-020-0181-9.

[26] C. Chen, Z. Zhang, Y. Cai, Y. Liu, and H. Chen, "Research and development status of *in situ* field assisted laser additive manufacturing: A review", *Optics & Laser Technology*, vol. 181, p. 111 700, Feb. 1, 2025, ISSN: 0030-3992. DOI: 10.1016/j.optlastec.2024.111700.

[27] S. Katayama, "Introduction: Fundamentals of laser welding*", in *Handbook of Laser Welding Technologies*, ser. Woodhead Publishing Series in Electronic and Optical Materials, S. Katayama, Ed., Woodhead Publishing, Jan. 1, 2013, pp. 3–16. DOI: 10.1533/9780857098771.1.3.

[28] A. Klimpel, "Review and Analysis of Modern Laser Beam Welding Processes", *Materials*, vol. 17, no. 18, p. 4657, Jan. 2024, ISSN: 1996-1944. DOI: 10.3390/ma17184657.

[29] A. Fysikopoulos, G. Pastras, J. Stavridis, P. Stavropoulos, and G. Chryssolouris, "On the Performance Evaluation of Remote Laser Welding Process: An Automotive Case Study", *Procedia CIRP*, Research and Innovation in Manufacturing: Key Enabling Technologies for the Factories of the Future - Proceedings of the 48th CIRP Conference on Manufacturing Systems, vol. 41, pp. 969–974, Jan. 1, 2016, ISSN: 2212-8271. DOI: 10.1016/j.procir.2016.01.005.

[30] E. Assunção, L. Quintino, and R. Miranda, "Comparative study of laser welding in tailor blanks for the automotive industry", *The International Journal of Advanced Manufacturing Technology*, vol. 49, no. 1–4, pp. 123–131, Jul. 2010, ISSN: 0268-3768, 1433-3015. DOI: 10.1007/s00170-009-2385-0.

[31] B. Chang et al., "Study of Gravity Effects on Titanium Laser Welding in the Vertical Position", *Materials*, vol. 10, no. 9, p. 1031, Sep. 2017, ISSN: 1996-1944. DOI: 10.3390/ma10091031.

[32] N. A. Steshenkova, N. A. Afanasiev, I. A. Tsibulskiy, and E. A. Valdaiytseva, "Application of system of geometrical and technological adaptation of process laser-arc hybrid welding for production in the minimum tolerances of hull designs of means of marine engineering", *Journal of Physics: Conference Series*, vol. 1109, p. 012 045, Nov. 2018, ISSN: 1742-6588, 1742-6596. DOI: 10.1088/1742-6596/1109/1/012045.

[33] J. T. Kahnamouei and M. Moallem, "Advancements in control systems and integration of artificial intelligence in welding robots: A review", *Ocean Engineering*, vol. 312, p. 119 294, Nov. 2024, ISSN: 00298018. DOI: 10.1016/j.oceaneng.2024.119294.

[34] Q. Guo et al., "Progress, challenges and trends on vision sensing technologies in automatic/intelligent robotic welding: State-of-the-art review", *Robotics and Computer-Integrated Manufacturing*, vol. 89, p. 102 767, Oct. 1, 2024, ISSN: 0736-5845. DOI: 10.1016/j.rcim.2024.102767.

[35] T. Adrien, S. Dehghan, A. Iltaf, and N. Barka, "A comprehensive study on laser welding of stainless steel 301LN in lap joint configuration", *Welding in the World*, vol. 68, no. 12, pp. 3141–3168, Dec. 1, 2024, ISSN: 1878-6669. DOI: 10.1007/s40194-024-01840-3.

[36] Y. Mi, F. Sikström, M. Nilsen, I. Choquet, and A. Ancona, "Adaptive laser beam shaping with a deformable mirror for gap bridging in autogenous butt joint welding", *Journal of Manufacturing Processes*, vol. 152, pp. 762–774, 2025, ISSN: 1526-6125. DOI: 10.1016/j.jmapro.2025.08.009.

[37] L. Quintino and E. Assunção, "Conduction laser welding", in *Handbook of Laser Welding Technologies*, Elsevier, 2013, pp. 139–162. DOI: 10.1533/9780857098771.1.139.

[38] K. Schwarzkopf, R. Rothfelder, M. Rasch, and M. Schmidt, "Two-Color-Thermography for Temperature Determination in Laser Beam Welding of Low-Melting Materials", *Sensors*, vol. 23, no. 10, p. 4908, Jan. 2023, ISSN: 1424-8220. DOI: 10.3390/s23104908.

[39] M. Courtois, M. Carin, P. L. Masson, S. Gaied, and M. Balabane, "A new approach to compute multi-reflections of laser beam in a keyhole for heat transfer and fluid flow modelling in laser welding", *Journal of Physics D: Applied Physics*, vol. 46, no. 50, p. 505 305, Nov. 2013, ISSN: 0022-3727. DOI: 10.1088/0022-3727/46/50/505305.

[40] L.-W. Hsu and A. Salminen, "Laser welding monitoring with multisensory data fusion: A brief review", *IOP Conference Series: Materials Science and Engineering*, vol. 1296, no. 1, p. 012 014, Dec. 1, 2023, ISSN: 1757-8981, 1757-899X. DOI: 10.1088/1757-899X/1296/1/012014.

[41] K. Schricker et al., "A comprehensive study on the influence of spatial power distribution on time-dependent keyhole behavior in laser beam welding of copper by means of high-speed synchrotron X-ray imaging", *Optics & Laser Technology*, vol. 192, p. 113 999, Dec. 1, 2025, ISSN: 0030-3992. DOI: 10.1016/j.optlastec.2025.113999.

[42] P. Yadav et al., "Reducing porosity and quality enhancement in keyhole laser welding under 3D configurable external magnetic field stimulation", *Materials & Design*, vol. 257, p. 114 461, Sep. 1, 2025, ISSN: 0264-1275. DOI: 10.1016/j.matdes.2025.114461.

[43] S. Shevchik et al., "Supervised deep learning for real-time quality monitoring of laser welding with X-ray radiographic guidance", *Scientific Reports*, vol. 10, no. 1, p. 3389, Feb. 25, 2020, ISSN: 2045-2322. DOI: 10.1038/s41598-020-60294-x.

[44] P. Jiang, C. Wang, Q. Zhou, X. Shao, L. Shu, and X. Li, "Optimization of laser welding process parameters of stainless steel 316L using FEM, Kriging and NSGA-II", *Advances in Engineering Software*, vol. 99, pp. 147–160, Sep. 1, 2016, ISSN: 0965-9978. DOI: 10.1016/j.advengsoft.2016.06.006.

[45] M. Haas et al., "Improving the weld seam quality in laser welding processes by means of Bayesian optimization", *Procedia CIRP*, 13th CIRP Conference on Photonic Technologies [LANE 2024], 15-19 September 2024, Fürth, Germany, vol. 124, pp. 772–775, Jan. 1, 2024, ISSN: 2212-8271. DOI: 10.1016/j.procir.2024.08.222.

[46] D. Wu et al., "Progress and perspectives of in-situ optical monitoring in laser beam welding: Sensing, characterization and modeling", *Journal of Manufacturing Processes*, vol. 75, pp. 767–791, Mar. 2022, ISSN: 15266125. DOI: 10.1016/j.jmapro.2022.01.044.

[47] Y. Kawahito and H. Wang, "In-situ observation of gap filling in laser butt welding", *Scripta Materialia*, vol. 154, pp. 73–77, Sep. 1, 2018, ISSN: 1359-6462. DOI: 10.1016/j.scriptamat.2018.05.033.

[48] F. Roland, L. Manzon, P. Kujala, M. Brede, and J. Weitzenbock, "Advanced Joining Techniques in European Shipbuilding", *Journal of Ship Production*, vol. 20, no. 03, pp. 200–210, Aug. 1, 2004, ISSN: 8756-1417. DOI: 10.5957/jsp.2004.20.3.200.

[49] Y.-W. Huang, X.-D. Gao, P. P. Gao, B. Ma, and Y.-X. Zhang, "Laser welding monitoring techniques based on optical diagnosis and artificial intelligence: A

review", *Advances in Manufacturing*, vol. 13, no. 2, pp. 337–361, Jun. 1, 2025, ISSN: 2195-3597. DOI: 10.1007/s40436-024-00493-1.

[50] B. Regaard, S. Kaierle, W. Schulz, and A. Moalem, "Advantages of coaxial external illumination for monitoring and control of laser materials processing", in *ICALEO 2005: 24th International Congress on Laser Materials Processing and Laser Microfabrication*, Miami, Florida, USA: Laser Institute of AmericaLIA, Oct. 1, 2005, p. 2307. DOI: 10.2351/1.5060481.

[51] P. Hildinger, T. Seefeld, and A. Bohlen, "Characterization of optical emissions during laser metal deposition for the implementation of an in-process powder stream monitoring", *Journal of Laser Applications*, vol. 35, no. 4, p. 042 048, Oct. 16, 2023, ISSN: 1042-346X. DOI: 10.2351/7.0001161.

[52] K. S. B. Ribeiro, H. H. L. Núñez, G. S. Venter, H. R. Doude, and R. T. Coelho, "A hybrid machine learning model for in-process estimation of printing distance in laser Directed Energy Deposition", *The International Journal of Advanced Manufacturing Technology*, vol. 127, no. 7–8, pp. 3183–3194, Aug. 2023, ISSN: 0268-3768, 1433-3015. DOI: 10.1007/s00170-023-11582-z.

[53] V. H. Mochi, H. H. L. Núñez, K. S. B. Ribeiro, and G. S. Venter, "Real-time prediction of deposited bead width in L-DED using semi-supervised transfer learning", *The International Journal of Advanced Manufacturing Technology*, vol. 129, no. 11–12, pp. 5643–5654, Dec. 2023, ISSN: 0268-3768, 1433-3015. DOI: 10.1007/s00170-023-12658-6.

[54] B. Boeckmans, M. Zhang, F. Welkenhuyzen, and J.-P. Kruth, "Determination of Aspect Ratio Limitations, Accuracy and Repeatability of a Laser Line Scanning CMM Probe", *International Journal of Automation Technology*, vol. 9, no. 5, pp. 466–472, Sep. 5, 2015. DOI: 10.20965/ijat.2015.p0466.

[55] H. Sun, L. Xia, Y. Zhou, W. Qian, W. Wang, and K. Zhang, "Online detection and evaluation of weld surface defects based on lightweight network VGG16-UNet and laser scanning", *Journal of Manufacturing Processes*, vol. 129, pp. 292–306, Nov. 15, 2024, ISSN: 1526-6125. DOI: 10.1016/j.jmapro.2024.08.037.

[56] V. Pasquinelli et al., "Laser line triangulation measurement on incandescent steel objects: Methodologies to improve optical signal to noise ratio", *Acta IMEKO*, vol. 14, no. 2, pp. 1–10, Jun. 25, 2025, ISSN: 2221-870X. DOI: 10.21014/actaimeko.v14i2.2069.

[57] M. Boley, F. Fetzer, R. Weber, and T. Graf, "Statistical evaluation method to determine the laser welding depth by optical coherence tomography", *Optics and Lasers in Engineering*, vol. 119, pp. 56–64, Aug. 1, 2019, ISSN: 0143-8166. DOI: 10.1016/j.optlaseng.2019.03.014.

[58] M. Sokolov, P. Franciosa, R. Al Botros, and D. Ceglarek, "Keyhole mapping to enable closed-loop weld penetration depth control for remote laser welding of aluminum components using optical coherence tomography", *Journal of Laser Applications*, vol. 32, no. 3, p. 032 004, Jun. 3, 2020, ISSN: 1042-346X. DOI: 10.2351/7.0000086.

[59] L.-W. Hsu, S. P. Chozaki, K. S. Barros Ribeiro, W. M. Bessa, and A. Salminen, "Evaluation of different surface topography measurement methodologies for characterizing weld beads in shipbuilding", *Procedia CIRP*, 13th CIRP Conference on Photonic Technologies [LANE 2024], 15-19 September 2024, Fürth, Germany, vol. 124, pp. 536–539, Jan. 1, 2024, ISSN: 2212-8271. DOI: 10.1016/j.procir.2024.08.169.

[60] J. Li, J. Zhao, Y. Luo, J. Huang, and Y. Ding, "Fusing optical coherence tomography and coaxial image for contact width and penetration depth monitoring

during laser lap welding", *Optics & Laser Technology*, vol. 183, p. 112 294, May 1, 2025, ISSN: 0030-3992. DOI: 10.1016/j.optlastec.2024.112294.

[61] C. Parikh et al., "Volumetric defect analysis in friction stir welding based on three dimensional reconstructed images", *Journal of Manufacturing Processes*, vol. 29, pp. 96–112, Oct. 1, 2017, ISSN: 1526-6125. DOI: 10.1016/j.jmapro.2017.07.006.

[62] C. Geiger, P. Garkusha, C. Bernauer, S. Mehrl, P. A. Schirmer, and M. F. Zaeh, "Acoustic process monitoring during the structuring of the diffusion media for fuel cells with Ultrashort Laser Pulses", *Procedia CIRP*, 13th CIRP Conference on Photonic Technologies [LANE 2024], 15-19 September 2024, Fürth, Germany, vol. 124, pp. 51–56, Jan. 1, 2024, ISSN: 2212-8271. DOI: 10.1016/j.procir.2024.08.069.

[63] N. Authier, E. Touzet, F. Lücking, R. Sommerhuber, V. Bruyère, and P. Namy, "Coupled membrane free optical microphone and optical coherence tomography keyhole measurements to setup welding laser parameters", in *High-Power Laser Materials Processing: Applications, Diagnostics, and Systems IX*, vol. 11273, SPIE, Mar. 2, 2020, pp. 41–52. DOI: 10.1117/12.2543999.

[64] L. Tomcic, A. Ederer, S. Grabmann, M. Kick, J. Kriegler, and M. F. Zaeh, "Interpreting acoustic emissions to determine the weld depth during laser beam welding", *Journal of Laser Applications*, vol. 34, no. 4, p. 042 052, Nov. 28, 2022, ISSN: 1042-346X. DOI: 10.2351/7.0000796.

[65] B. Fischer, "Optical microphone hears ultrasound", *Nature Photonics*, vol. 10, no. 6, pp. 356–358, Jun. 2016, ISSN: 1749-4893. DOI: 10.1038/nphoton.2016.95.

[66] Q. G. Ma, X. L. Sun, H. J. Fu, D. C. Zhao, and J. F. Guo, "Manufacturing Process Design Based on Mental and Physical Workload Analysis", *Applied*

Mechanics and Materials, vol. 345, pp. 482–485, 2013, ISSN: 1662-7482. DOI: 10.4028/www.scientific.net/AMM.345.482.

[67] J. V. Kresta, J. F. Macgregor, and T. E. Marlin, "Multivariate statistical monitoring of process operating performance", *The Canadian Journal of Chemical Engineering*, vol. 69, no. 1, pp. 35–47, 1991, ISSN: 1939-019X. DOI: 10.1002/cjce.5450690105.

[68] P. M. Ramos, R. C. Martins, S. Rapuano, and P. Daponte, "Frequency and Time—Frequency Domain Analysis Tools in Measurement", in *Data Modeling for Metrology and Testing in Measurement Science*, F. Pavese and A. B. Forbes, Eds., Boston: Birkhäuser, 2009, pp. 1–27. DOI: 10.1007/978-0-8176-4804-6_6.

[69] Z. Kunpeng, W. Y. San, and H. G. Soon, "4 - Signal processing for tool condition monitoring: From wavelet analysis to sparse decomposition", in *Mechatronics and Manufacturing Engineering*, ser. Woodhead Publishing Reviews: Mechanical Engineering Series, J. P. Davim, Ed., Woodhead Publishing, Jan. 1, 2012, pp. 115–157. DOI: 10.1533/9780857095893.115.

[70] W. Y. Yoon and N. K. Kwon, "Integrated Neural Network Approach for Enhanced Vital Signal Analysis Using CW Radar", *Electronics*, vol. 13, no. 13, p. 2666, Jan. 2024, ISSN: 2079-9292. DOI: 10.3390/electronics13132666.

[71] H. Li et al., "Machine learning for high-accuracy signal processing in ultra cryogenic temperature optical fiber sensors", *Optics Express*, vol. 33, no. 16, pp. 33 216–33 231, Aug. 11, 2025, ISSN: 1094-4087. DOI: 10.1364/OE.568626.

[72] W. Huang and R. Kovacevic, "Acoustic monitoring of weld penetration during laser welding of high strength steels", presented at the ICALEO 2009: 28th International Congress on Laser Materials Processing, Laser Microprocessing

and Nanomanufacturing, AIP Publishing, Nov. 1, 2009, pp. 630–637. DOI: 10.2351/1.5061620.

[73] L. Liu, B. Shen, and X. Wang, "Research on Kernel Function of Support Vector Machine", in *Advanced Technologies, Embedded and Multimedia for Human-centric Computing*, Y.-M. Huang, H.-C. Chao, D.-J. Deng, and J. J. (H. Park, Eds., Dordrecht: Springer Netherlands, 2014, pp. 827–834. DOI: 10.1007/978-94-007-7262-5_93.

[74] I. K. Pious, A. Rajalakshmi, P. K. R, V. C M, M. Nalini, and S. S. R, "Enhancing Prediction Accuracy Through Random Forest in Classification and Regression", in *2024 International Conference on Smart Technologies for Sustainable Development Goals (ICSTSDG)*, Nov. 2024, pp. 1–6. DOI: 10.1109/ICSTSDG61998.2024.11026358.

[75] I. P. Thatikonda, S. Jannu, C. Thuppari, G. Kalyani, and M. Swathi, "An Advanced Heart Disease Prediction Model with Gradient Boosting: A Data-Driven Approach", in *2025 6th International Conference on Recent Advances in Information Technology (RAIT)*, Mar. 2025, pp. 1–6. DOI: 10.1109/RAIT65068.2025.11089262.

[76] G. Maculotti, G. Genta, and M. Galetto, "Optimisation of laser welding of deep drawing steel for automotive applications by Machine Learning: A comparison of different techniques", *Quality and Reliability Engineering International*, vol. 40, no. 1, pp. 202–219, 2024, ISSN: 1099-1638. DOI: 10.1002/qre.3377.

[77] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks", *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 24, 2017, ISSN: 0001-0782. DOI: 10.1145/3065386.

[78] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[79] K. Simonyan and A. Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition". arXiv: 1409.1556 [cs], pre-published.

[80] O. Ronneberger, P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". arXiv: 1505.04597 [cs], pre-published.

[81] A. Dosovitskiy et al. "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale". arXiv: 2010.11929 [cs], pre-published.

[82] D. P. Kingma and M. Welling, "An Introduction to Variational Autoencoders", *Foundations and Trends® in Machine Learning*, vol. 12, no. 4, pp. 307–392, 2019, ISSN: 1935-8237, 1935-8245. DOI: 10.1561/2200000056. arXiv: 1906.02691 [cs].

[83] Q. Khan, "Transfer Learning for Computer Vision: From Theory to Implementation", in *Advances in Computational Intelligence and Robotics*, M. I. Uddin and W. K. Mashwani, Eds., IGI Global, Feb. 23, 2024, pp. 26–40. DOI: 10.4018/979-8-3693-1738-9.ch002.

[84] O. Elharrouss, Y. Akbari, N. Almadeed, and S. Al-Maadeed, "Backbones-review: Feature extractor networks for deep learning and deep reinforcement learning approaches in computer vision", *Computer Science Review*, vol. 53, p. 100 645, Aug. 1, 2024, ISSN: 1574-0137. DOI: 10.1016/j.cosrev.2024.100645.

[85] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge", *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, Dec. 1, 2015, ISSN: 1573-1405. DOI: 10.1007/s11263-015-0816-y.

[86] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python", *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011, ISSN: 1533-7928.

[87] M. Barandas et al., "TSFEL: Time Series Feature Extraction Library", *SoftwareX*, vol. 11, p. 100 456, Jan. 1, 2020, ISSN: 2352-7110. DOI: 10.1016/j.softx.2020.100456.

[88] P. Virtanen et al., "SciPy 1.0: Fundamental algorithms for scientific computing in Python", *Nature Methods*, vol. 17, no. 3, pp. 261–272, Mar. 2020, ISSN: 1548-7105. DOI: 10.1038/s41592-019-0686-2.

[89] A. Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". arXiv: 1912.01703 [cs], pre-published.