

MASTER'S THESIS

Robust Anomaly Detection With Voltage Data of Wire-Arc Additive Manufacturing

Robuste Anomalieerkennung in Spannungsdaten von Wire-Arc Additive Manufacturing

In partial fulfillment of the requirements for the degree

Master of Science of RWTH Aachen University

at the

Chair of Information Management in Mechanical Engineering (IMA)

Submitted by:	Anshul pagariya
Matriculation No.:	404634
Field of Study:	MSc Robotics system engineering
Supervisor:	Prof. Dr. Phil. Ingrid Isenhardt, MSc Hans zhou
External advisor:	Dr. Philip clissold howell
Examiner:	Prof. Dr. Phil. Ingrid Isenhardt
Place:	Aachen
Date:	November 22, 2022

Disclaimer

I confirm that this Master's Thesis is my own work and I have documented all sources and material used.

Aachen, 21 November, 2022

A handwritten signature in blue ink, appearing to be 'Anshul', is written on a yellow rectangular background.

(Anshul pagariya)

Abstract

The thesis developed different approaches based on features and raw data of voltage signals from Wire arc additive manufacturing (WAAM). Anomaly detection models were able to detect anomalies in voltage data during oxidation, near pores and large cavities indicating that detected anomalies led to defect formation, although model is not reliable for detecting formation of pores and cavities. Relevant features and their relationship with defects were also discussed, which gives more insight into data. Analysis of real world factory data and discussion on transferability of models developed for lab data to factory data was also done. Work done sets up a base on which further research could be done to robustly detect anomalies in WAAM via voltage data and finally use developed model for mitigation strategies.

Zusammenfassung

In dieser Arbeit wurden verschiedene Ansätze auf der Grundlage von Merkmalen und Rohdaten der Spannungssignale von Wire arc additive manufacturing (WAAM) entwickelt. Die Modelle zur Erkennung von Anomalien waren in der Lage, Anomalien in den Spannungsdaten während der Oxidation, in der Nähe von Poren und großen Hohlräumen zu erkennen, was darauf hindeutet, dass die erkannten Anomalien zur Defektbildung führten, obwohl das Modell nicht zuverlässig für die Erkennung der Bildung von Poren und Hohlräumen ist. Relevante Merkmale und ihre Beziehung zu Defekten wurden ebenfalls erörtert, was einen besseren Einblick in die Daten ermöglicht. Außerdem wurden reale Fabrikdaten analysiert und die Übertragbarkeit der für Labordaten entwickelten Modelle auf Fabrikdaten diskutiert. Die geleistete Arbeit bildet eine Grundlage, auf der weitere Forschungen zur robusten Erkennung von Anomalien in WAAM anhand von Spannungsdaten und schließlich zur Verwendung des entwickelten Modells für Abhilfestrategien durchgeführt werden können.

I Contents

I	Contents	4
li	Abbreviations	7
lii	List of Figures	9
iv	List of Tables	12
1	Introduction	13
1.1	Motivation of Research	13
1.2	Challenges of Research.....	13
2	Background and Related Work	15
2.1	Wire Arc Additive Manufacturing.....	15
2.1.1	Introduction to Cmt.....	17
2.1.2	Defects in Waam.....	19
2.2	In Situ Monitoring.....	20
2.3	Machine Learning	23
2.3.1	Neural Network	24
2.3.2	Bayesian Optimization.....	28
2.4	Anomaly Detection.....	29
2.5	Dimensionality Reduction.....	30
2.6	3d Coordinate Reference Changes.....	30
2.7	Algorithms.....	31
2.7.1	Convolutional Neural Network.....	31
2.7.2	Autoencoder.....	33
2.7.3	Inception Network.....	34
3	Experimental Setup, Data Creation and Processing	36
3.1	Lab and Factory Setup.....	36
3.1.1	Lab Setup and Sensor.....	36
3.1.2	Factory Setup and Sensor.....	37
3.2	Different Kinds of Data	37
3.2.1	Linear Wall (Wall No. 1).....	38
3.2.2	Oxidized Complex Wall (Wall No. 2).....	39
3.2.3	Inverted Pyramid Complex Wall (Wall No. 3).....	40
3.2.4	Airfoil Blades (Wall No. 4).....	40
3.3	Data Processing and Labeling	41
3.3.1	Data Cleaning and Preprocessing.....	41
3.3.2	Labeling	41
3.4	Initial Data Analysis.....	43
3.5	Training Data Creation	44
4	Methodology for Lab Data	45
4.1	Model Evaluation Algorithm.....	45
4.2	Feature Engineering	46
4.3	Model Development Using Supervised Learning.....	47
4.4	Model Development Based on Forecasting Using Unsupervised Learning.....	48
4.5	Model Development Based on Single Cmt Cycle Using Unsupervised Learning...	48
4.5.1	Reconstruction-Based Learning	49
4.5.2	Feature-Based Learning.....	50
4.6	Model Development Based on Multiple Cmt Cycles Using Unsupervised Learning	50
5	Methodology for Factory Data	53

6	Results and Discussion	55
6.1	Result for Model Based on Forecasting In Unsupervised Approach	55
6.2	Result for Model Based on Single Cmt Cycle Using Unsupervised Learning.....	55
6.2.1	Reconstruction-Based Learning	55
6.2.2	Feature-Based Learning.....	59
6.3	Result for Model Based on Multiple Cmt Cycle Using Unsupervised Learning	62
6.4	Transferability on Factory Data	66
7	Conclusions and Future Work.....	67
7.1	Conclusion	67
7.2	Future Work	67
8	Bibliography.....	67

Abbreviations

Abbreviations	Description
WAAM	Wire Arc Additive Manufacturing
GMAW	Gas Metal Arc Welding
GTAW	Gas Tungsten Arc Welding
CT	Computerized Tomography
CMT	Cold Metal Transfer
DNN	Deep Neural Network
NN	Neural Network
CNN	Convolutional Neural Network
AM	Additive Manufacturing
MIG	Metal Inert Gas
MAG	Metal Active Gas
PCA	Principal Component Analysis
SGD	Stochastic Gradient Descent

II List of Figures

2.1	State of additive materials for 3D metal printing [Li + 19].....	15
2.2	Setup of welding torch setup for layer-by-layer part production [Hau + 21a].....	16
2.3	Turbine blades manufactured using WAAM [Fel +19].....	16
2.4	Cold metal transfer process [Wes + 10].....	18
2.5	Current and Voltage waveform of CMT process [Sel + 18].....	18
2.6	Example of porosity in WAAM [Hau + 21b].....	19
2.7	Example of oxidation in WAAM from part printed in siemens lab.....	20
2.8	Example of large cavities in WAAM.....	20
2.9	Schematic from iMust, a monitoring system [Xu + 18].....	21
2.10	Schematic from monitoring system [Rei + 20a].....	22
2.11	Comparison of voltage waveform with and without porosity occurrence [Shi + 20].....	22
2.12	Neural network [SV14].....	25
2.13	Example of gaussian process and bayesian optimization [Ngu + 19].....	29
2.14	Rotation of 0-xyz by an angle α about axis z [BLL10].....	31
2.15	Local receptive field of first hidden neuron. [M16].....	32
2.16	DeepAnt architecture proposed by Munir et al. [Mun + 19].....	32
2.17	Convolutional autoencoder architecture for anomaly detection [Tan + 19].....	33
2.18	Example of autoencoder. [Ban + 20]	33
2.19	Inception module in Inception network [Isn + 20].....	34
2.20	Inception network with its componenets [Isn + 20].....	35
3.1	WAAM system cell. [Rei + 20b].....	37
3.2	Linear wall and its x-ray (left), printed in lab.....	38
3.3	Voltage and current waveform (wall no. 1).....	38
3.4	Robot path during meander	39
3.5	CT scan and part view (wall no. 2).....	39
3.6	Voltage and current waveform (wall no. 3 and 4).....	39
3.7	CT scan and part view (wall no. 3).....	40
3.8	Robot path for printing airfoil for a particular layer.....	40
3.9	Voltage and current waveform (wall no. 4).....	41
3.10	CT scan analysis. [myv]	42
3.11	XY cross-sectional view wall no. 2	43
3.12	XY cross-sectional view wall no. 3.....	43
3.13	Boxplots of standard deviations for all layers.....	44
3.14	A cross-sectional view of layer number 19	44
4.1	Instance of current and voltage CMT cycles during a) normal b) low gas flow region (oxidized).....	46
4.2	Voltage CMT cycle's phases segmentation using developed phase transition thresholds based algorithm.....	46
4.3	Forecasted and original signal prediction by Inception network.....	48
4.4	Architecture of reconstruction-based convolutional network.....	49
4.5	Instance of original input and reconstruction by developed convolutional network....	49
4.6	Architecture of feature-based autoencoder.....	50
4.7	Boxplots of feature: average of maximum value of 6 consecutive voltage CMT cycles for each layer.....	51

4.8	Boxplots of feature: average of mean background phase of 6 consecutive voltage CMT cycles for each layer.....	51
4.9	Boxplot of feature: variations of background phase of 6 consecutive voltage CMT cycles for each layer.....	52
4.10	Architecture of developed autoencoder.....	52
5.1	Voltage cycle of factory and lab.....	53
5.2	Current cycle of lab and factory data.....	53
5.3	Filtered voltage data	54
6.1	Layer no. 5th's XY cross-sectional view of CT scan with low volume defects (colored) overlaid with model's results.....	56
6.2	Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 5th.....	56
6.3	Anomalous and normal data detected by model in layer 11 with low gas flow (oxidized) overlaid over a CT scan picture of layer.....	57
6.4	Layer no. 13th XY cross-sectional view of CT scan with large defects (colored) overlaid with model's results.....	58
6.5	Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 13th.....	58
6.6	Layer no. 12th XY cross-sectional view of CT scan with large defects (colored) overlaid with model's results.....	58
6.7	Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 12.....	59
6.8	Layer no. 5th's XY cross sectional view of CT scan with low volume defects (colored) overlaid with model 's results.....	60
6.9	Anomalies and normal data detected by model in layer 11 with low gas flow (oxidized) overlaid over CT scan of layer.....	60
6.10	Layer no. 13th XY cross-sectional view of CT scan with large defects (colored) overlaid with model's results.....	61
6.11	Layer no. 12th XY cross-sectional view of CT scan with large and small defects (colored) overlaid with model's results.....	61
6.12	Layer no. 5th's XY cross sectional view of CT scan with low volume defects (colored) overlaid with model 's results.....	62
6.13	Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 5th.....	63
6.14	Anomalies and normal data detected by model in layer 11 with low gas flow (oxidized) overlaid over CT scan of layer.	63
6.15	Principal components of voltage data.....	63
6.16	Layer no. 13th XY cross-sectional view of CT scan with large defects (colored) overlaid with model's results.	64
6.17	Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 13th.	65
6.18	Layer no. 12th XY cross-sectional view of CT scan with large defects (colored) overlaid with model 's results.	65
6.19	Correlation found between model anomalies and defect (joined by lines) by using evaluation algorithm for layer 12.....	66

III List of Tables

2.1	Feature extracted for voltage data [Shi + 20].....	23
2.2	Confusion matrix	27
3.1	Sensor details (lab).....	36
3.2	Sensor details (factory).....	37
3.3	Different kinds of data.....	38
4.1	Voltage CMT cycle features and descriptions.....	47
6.1	Quantitative results for wall no. 2 with reconstruction-based model using model evaluation algorithm.....	56
6.2	Quantitative results for wall no. 3 with reconstruction-based model using model evaluation algorithm.....	57
6.3	Quantitative results for wall no. 2 with feature-based model using model evaluation algorithm.	59
6.4	Quantitative results for wall no. 3 with feature-based model using model evaluation algorithm.....	60
6.5	Quantitative results for wall no. 2 with trained model using model evaluation algorithm.	62
6.6	Quantitative results for wall no. 3 with trained model using model evaluation algorithm.	64

1 Introduction

Humanity's desire to push its boundaries and improve its livelihoods has driven many technological advances. We desire to produce objects which help us in our day-to-day activities or to transport goods fast; with each day passing by we are improving efficiency and scope of technologies. In recent times interest in additive manufacturing (AM) has increased greatly due to its cost-effectiveness, efficiency, fast-to-use, and complex production capabilities. To make large-scale use of AM for production possible its reliability, repeatability, and product quality needs to be improved.

With advancement in industry 4.0, machines are interconnected, making real-time process parameters monitoring possible, this naturally has motivated researchers to monitor printing process of additive manufacturing techniques. This becomes especially important in case of additive manufacturing where part is built layer by layer, detecting and understanding anomalies in real time helps significantly improve part quality.

1.1 Motivation of Research

Interest in additive manufacturing has grown recently as a result of numerous novel technologies and expanding potential applications. Welding-based Wire Arc Additive Manufacturing (WAAM) process, which is a direct energy deposition process in additive manufacturing is currently used in a variety of industrial sectors including automotive, aviation, rail, and medical engineering industries.

Main issue with industrialization of WAAM is lack of quality and repeatability of manufactured part. While building a part through WAAM several anomalies could occur like pores, oxidation of metal, discontinuities, etc., which if undetected during part build-up process could damage or build a part of bad quality. One solution is to detect anomalies, which arise during WAAM printing process in real-time and notify a machine operator for adjusting process parameters. Detection of anomalies requires to be in real-time, thus sensors measuring e.g., voltage and current are used to monitor whole printing process. This thesis uses voltage data acquired from Siemens research lab in Munich and Siemens factory (production) to develop a anomaly detection system.

1.2 Challenges of Research

Given requirements set by objective, development of a detection system for defect inducing anomalies involves several challenges:

- **Imbalanced data** - Anomalies are rare and therefore training algorithms to learn the representations of anomalous data becomes difficult.

- **Unprecise labeling** – For labeling, spatial anomalies in printed part need to be correlated with the time series data of voltage. Because of lack of clear separation between layers of the printed part and other problems like pores traveling during the printing process, it becomes difficult to find point of origin, and hence labeling process is inaccurate and unprecise.
- **Low false negative rate-** It is important that anomaly detection system developed has low false negatives because a missed anomaly can lead to part failure.
- **Real-time anomaly detection** - Anomaly detection system should be able to detect anomalies in real-time during printing process to give an alarm of nonoptimal process parameters or abnormalities in printed part to operator. These characteristics constrain time required for real-time data analysis of sensor data.

2 Background and Related Work

In this chapter fundamentals of WAAM, in-situ monitoring literature review, machine learning, anomaly detection fundamentals, and other methods and techniques used in thesis are discussed.

2.1 Wire Arc Additive Manufacturing

Over the last two decades, additive manufacturing (AM) has seen rapid industrial development. Previously, it was referred to as rapid prototyping (RP), rapid tooling (RT), and layered manufacturing (LM) [Li + 19]. AM technology can be classified according to the material and technology used. Figure 2.1 shows a classification based on material.

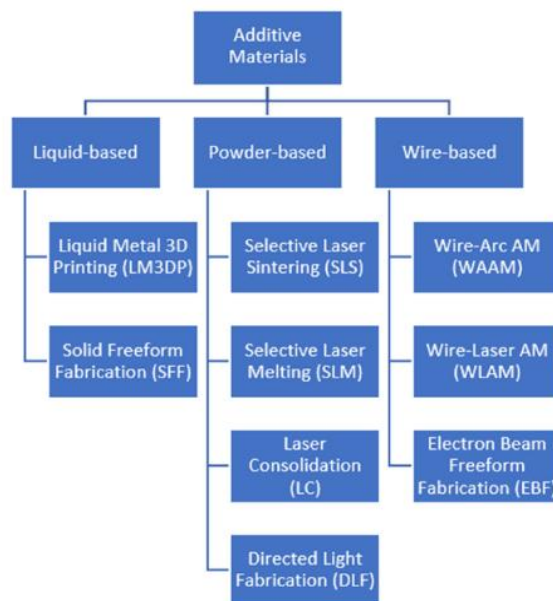


Figure 2.1: State of additive materials for 3D metal printing [Li + 19].

There are three different groups in wire based additive manufacturing (WBAM): laser-based, arc welding-based, and electron beam-based. Three of these groups use same process but have different power sources and deposition rates. [Li + 19] WAAM is based on the fundamental concepts of automated welding processes like gas metal arc welding (GMAW), plasma arc welding (PAW), and gas tungsten arc welding (GTAW). Most common arc welding process in WAAM is GMAW, also known as metal inert gas (MIG)/metal active gas (MAG). GMAW is fusion based arc welding. In GMAW arc is created between the tip of a feedstock wire and the workpiece. Weldpool and adjacent material are shielded by inert or active shielding gas. GMAW deposition rates range from 15 to 160 g/min, depending on the deposited

material and process parameters, making it ideal for producing large-scale parts in short time spans. [Rod +19] WAAM process is schematically shown in the figure 2.2.

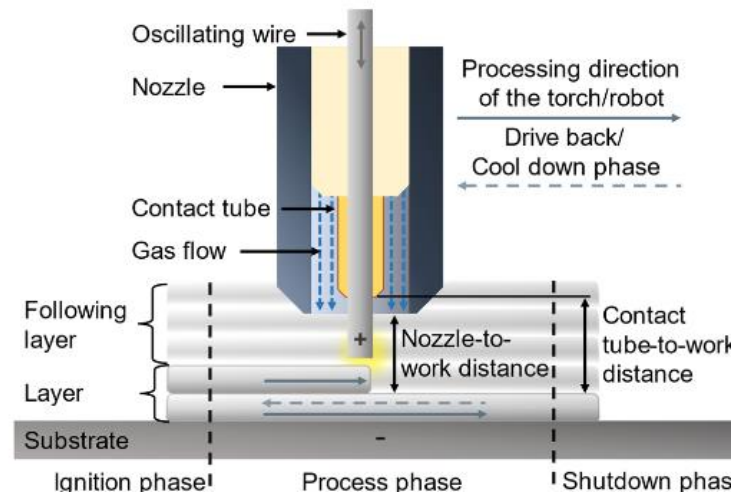


Figure 2.2: Schematic setup of welding torch setup for layer-by-layer part production [Hau + 21a].

A piece is constructed by first fusing first layer to a metallic surface and then depositing a new layer over the previous one. A layer is deposited by melting a portion of the previous layer, resulting in a pool of molten metal known as melt pool. Metal droplets from the feedstock wire are then added to melt pool iteratively while welding torch moves. There is no more heat input in that section as torch moves away, and the deposited metal solidifies into part of new layer. During operation, welding torch emits shielding gas, which prevents contamination of process from the air atmosphere, which would otherwise result in defects such as oxidation. A part manufactured by WAAM is shown in the figure 2.3.



Figure 2.3: Turbine blades manufactured using WAAM [Fel +19].

Manufacturing Ti6Al4V parts with WAAM reduced costs by 20-45% compared to electron beam additive manufacturing and 69-79% compared to direct metal laser sintering. [Cun + 17]

Benefits of WAAM :

1. **Parts size** - Maximum part size possible is as large as manipulator's reach capability.
2. **Material cost and utilization** - Wire is a low-cost feedstock that avoids many of the challenges associated with powder, such as particle size and distribution control, which affects process performance. Finally, unlike powder, wire is completely molten at point of deposition and becomes part of final structure, with a low likelihood of contamination.
3. **Deposition rates** - It is possible to achieve a high deposition rate. Most parts can be manufactured in one working day, with rates ranging from 1 kg/h to 4 kg/h for aluminum and steel, respectively. [Wil + 16]

To ensure a good final part quality in WAAM is difficult, because several process parameters influence quality of part. Voltage, current, shielding gas type and gas flow, contact to work distance, torch angle, travel speed, and wire feed speed all influence quality of the part. [Rod +19] A variant of GMAW known as cold metal transfer (CMT) has been adapted to additive manufacturing in pursuit of a better and more stable process to control molten metal deposition with reduced heat input. CMT procedure is covered in the following section.

2.1.1 Introduction to CMT

Recent advancements in metal additive manufacturing and new arc processes such as CMT process aided in gaining widespread attention. Fronius introduced CMT process, an innovative GMAW process, in 2004. [Pla + 19] It has digital process control and a special weld torch with an integrated servomotor that allows wire to be withdrawn every time it comes into contact with melt pool. This movement mechanically supports transfer of metal droplet, which aids in controlled molten metal deposition and thus reduces heat input. The process has several advantages, including increased process stability, stable arc length, less residual stress, and reduced dilution [Pla + 19]. During metal transfer, current drops to near zero, preventing spatter generation. [Sel + 18] These characteristics make it highly suitable for WAAM process and it is therefore widely used in research projects and commercially. Schematic diagram of droplet transfer and wire movement is shown in the figure 2.4. Fronius also developed cold metal transfer variants such as CMT pulse (CMT-P), CMT advanced (CMT-ADV), and CMT pulse advanced (CMT-PADV). [Rod +19]

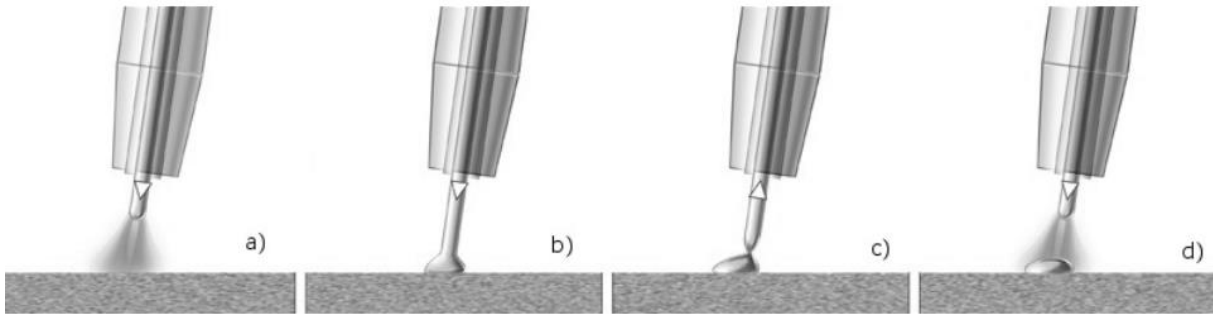


Figure 2.4: Cold metal transfer process **a)** arc ignition **b)** short circuit **c)** inversion of wire feed direction **d)** arc reignition [Wes + 10]

A typical CMT welding electrical signal cycle is the period required to deposit a molten electrode droplet into weld pool. Study of the energy distribution of different phases in droplet transfer process requires the analysis of current and voltage waveforms shown in the figure 2.5. The cycle is divided into three stages, which are as follows:

- **Peak current phase:** This phase corresponds to a high current pulse that easily ignites welding arc and then heats wire electrode to form a droplet.
- **Background current phase:** In this phase to prevent the globular transfer of small liquid droplet formed on wire tip, current is reduced. During this phase, wire moves towards the molten pool. This phase will continue until a short circuit occurs.
- **Short-circuiting phase:** Arc voltage is reduced to zero during this phase. At the same time, return signal is sent to wire feeder, which causes wire to draw back. This phase aids in liquid fracture and material transfer into the welding pool. [Sel + 18]

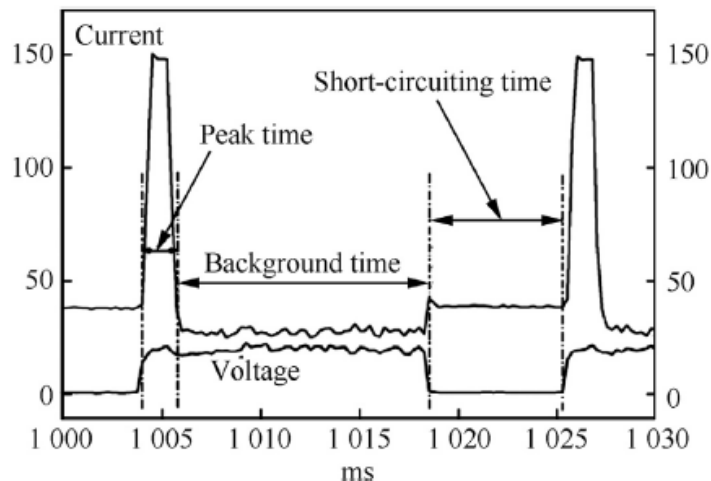


Figure 2.5: Current and voltage waveform of CMT process [Sel + 18].

2.1.2 Defects in WAAM

Since WAAM process is fundamentally similar to welding, defects such as hot cracking, cold cracking, porosity, delamination, and spatter for various alloys are well documented.[Rod +19] WAAM defects can occur for a variety of reasons, including poor programming strategy, unstable weld pool dynamics due to poor parameter setup, thermal deformation caused by heat accumulation, environmental influence (such as gas contamination), and other machine malfunctions. [Wu + 18] During deposition of one layer in WAAM, there is also a risk of defect formation and propagation to subsequent layers. Defects considered in this work are explained in the following paragraph.

Porosity: Porosity is a small gap or cavity in the metal. Porosity reduces the overall density of the parts and thus reduces the mechanical properties (such as tensile strength) of the manufactured components, which is a major challenge in welding applications. [Jin + 20; Hau + 21b]

Reasons for porosity in WAAM are:

- **Raw material based:** WAAM raw materials, including as-received wire and substrate, frequently have surface contamination, such as moisture, grease, and other hydrocarbon compounds, which can be difficult to remove completely. These contaminants can easily be absorbed into molten pool and cause porosity after solidification.
- **Process-induced:** Process-induced porosity is typically non-spherical and is caused primarily by insufficient path planning or an unstable deposition process. Insufficient fusion or spatter ejection is easily produced when deposition path is complex or manufacturing process is changeable, resulting in gaps or voids in these influenced regions. [Wu + 18]

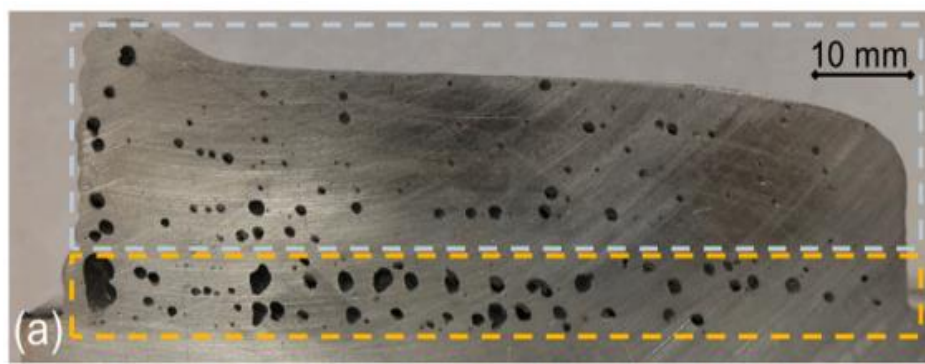


Figure 2.6: Example of porosity in WAAM from [Hau + 21b].

Oxidation: Shielding gas's role in welding applications is to prevent contact between melt pool and chemicals in the air: hydrogen, oxygen, and nitrogen. Melt pool may not be effectively shielded if the gas flow is insufficient or is affected by air currents. The figure 2.7 shows the

oxidized regions formed when gas flow was drastically reduced in WAAM for steel alloy in experiments conducted in Siemens lab, Munich.

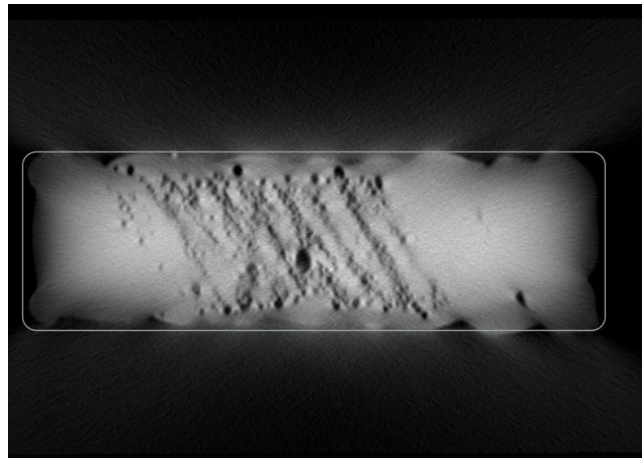


Figure 2.7: Example of oxidation in WAAM from part printed in siemens lab.

Discontinuity: These are large unwanted gaps formed between welding tracks as a result of deposition in an undesirable location or form deviation in an already printed part. When the next layer goes over the spot where there is a discontinuity, discontinuity will likely repeat itself, resulting in a cavity passing through multiple layers. Figure 2.8 shows cross-sectional CT scan view of printed part in Siemens lab, Munich.

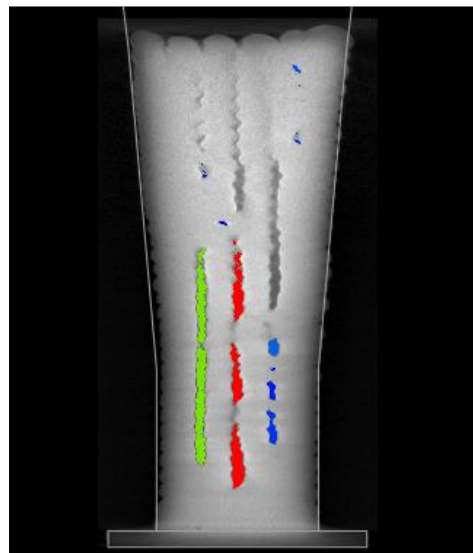


Figure 2.8: Example of large cavities in WAAM.

2.2 In-Situ Monitoring

Monitoring part build-up assists in understanding link between monitored signals and their effect on the component, which aids in development of a closed-loop control system. [Rod +19]

In WAAM final output is dependent on process parameters and non-destructive testing after the part is produced is expensive. [Xu + 18]

Many researchers have attempted to solve problem of WAAM printing defects. Sensors must be used to inspect complex parts while they are being printed, during fusion welding, the part is traditionally monitored by measuring voltage, shielding gas flow rate, travel speed, and wire feed speed. Sensors can also be used during WAAM to monitor the temperature in different regions, measure the size and geometry of beads, determine weld pool characteristics, monitor acoustic signals generated during deposition, detect electrical conductivity variations, and measure oxygen levels. [Rod +19]

Xu et al. [Xu + 18] proposed iMust model to monitor WAAM process. iMust is divided into 5 layers.

1. **Sensor layer:** All physical devices are housed in this layer.
2. **Signal layer:** This layer gathers raw signals from sensors.
3. **Feature layer:** A data processing procedure is used in this layer to extract information from raw data coming from Signal layers.
4. **Decision-making layers:** This layer will adjust process parameters to improve deposition quality.
5. **Output:** Output will be the recommended process parameters.

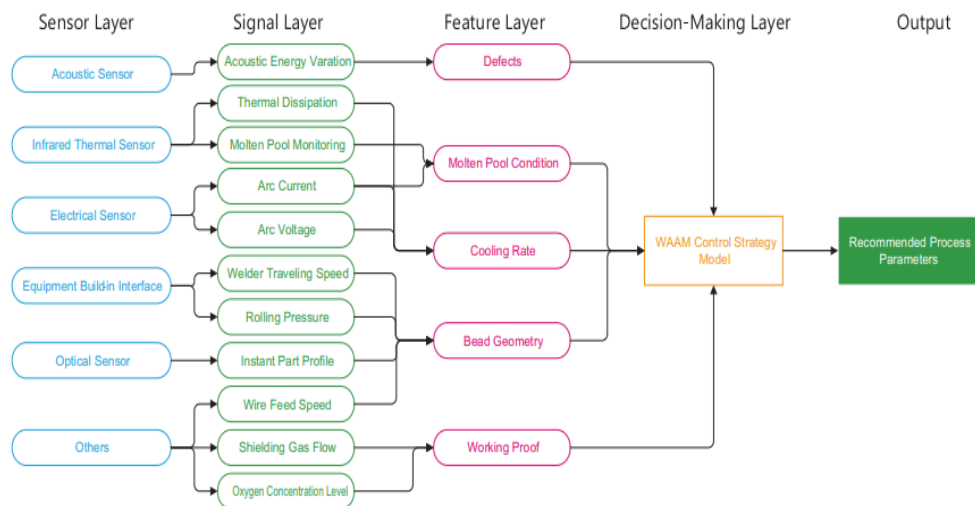


Figure 2.9: Schematic from iMust, a monitoring system proposed by Xu et al. [Xu + 18]

Reisch et al. [Rei + 20a] used separate deep neural network (DNN) models for each voltage, current, and Image data to predict the next step in time. Error between predicted and measured data for each sensor data was compared with the statistical distribution of historical errors to detect anomalous data. Mahalanobis distance was used to compute the likelihood of anomaly. Historical errors were computed for a sliding window of data, and a gradual increase in the error distance due to concept drift had relatively minor effects on anomaly detection. The figure 2.10 shows steps in monitoring system developed by Reisch et al. Aluminum alloy was used

as printing material. The approach was tested on industrial WAAM data the result showed that the approach was able to detect anomalies like oxidation, form deviation, and polluted surface.

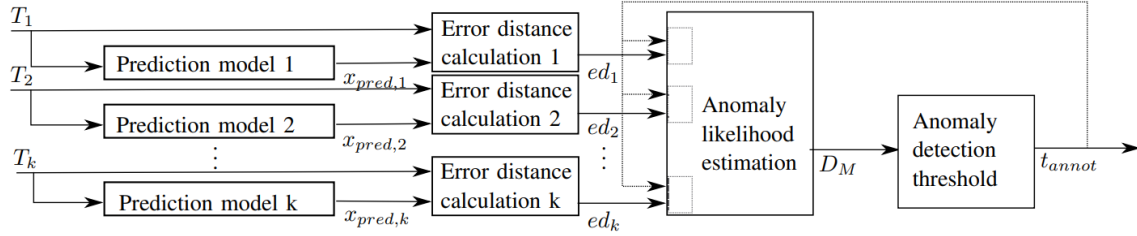


Figure 1. Approach of the proposed multivariate anomaly detector

Figure 2.10: Schematic from monitoring system proposed by Reisch et al. [Rei + 20a]

Huang et al. [Hua + 20] implemented monitoring and control system for aluminum alloys to control porosity. They started by extracting features from welding spectral data. A neural network was used to predict porosity based on the features. Finally, a controller based on Fuzzy-PID was developed to regulate and control porosity defect by adjusting welding current. Shin et al. [Shi + 20] found correlation between porosity and voltage waveform, shown in the figure 2.11, they extracted some features (table 2.1) from the chunk of 0.1 seconds of voltage data, and finally trained a deep neural network (DNN) and an artificial neural network to classify whether the 0.1 seconds of voltage feature data will cause a defect or not. Models were trained in a supervised fashion by gathering labeled data for porosity. Trained DNN achieved porosity prediction accuracy of 90%. Steel alloy was used as welding material.

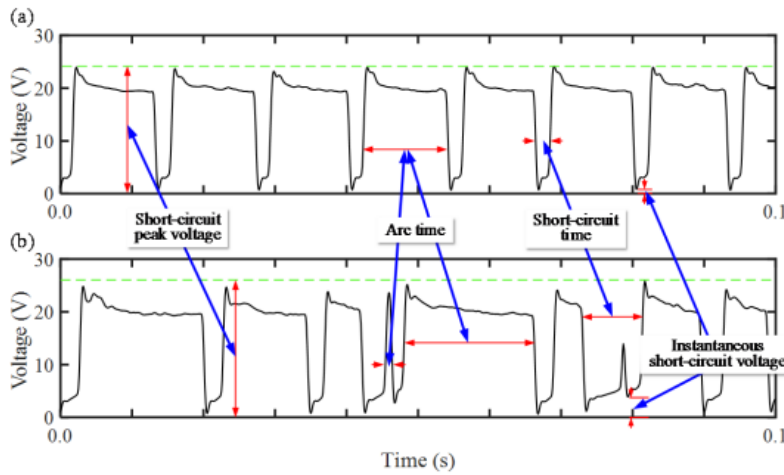


Figure 2.11: Comparison of voltage waveform with and without porosity occurrence a) no porosity signal, b) porosity signal [Shi + 20]

Feature Variable	Description	Symbol
X_1	Standard deviation of voltage	$s[V]$
X_2	Standard deviation of instantaneous short-circuit voltage	$s[V_s]$
X_3	Standard deviation of short-circuit peak voltage	$s[V_p]$
X_4	Average voltage during short-circuit time	$\bar{V}(T_s)$
X_5	Average voltage during arc time	$\bar{V}(T_a)$
X_6	Average short-circuit time	\bar{T}_s
X_7	Average arc time	\bar{T}_a
X_8	Number of short-circuit periods	$N[T]$
X_9	Standard deviation of short-circuit time	$s[T_s]$
X_{10}	Standard deviation of arc time	$s[T_a]$
X_{11}	Standard deviation of voltage during short-circuit time	$s[V(T_s)]$
X_{12}	Standard deviation of voltage during arc time	$s[V(T_a)]$

Table 2.1: Feature extracted for voltage data by Shin et al. [Shi + 20]

2.3 Machine Learning

Machine learning can be used to learn complex patterns in data. In this thesis, machine learning models were used to develop anomaly detection models. Machine learning is the task of creating a computer program that automatically improves with experience. In traditional engineering, design flow begins with acquiring domain knowledge which yields a mathematical model that captures physics of the underlying system. In contrast, machine learning approach substitutes the task of acquiring domain knowledge with easier task of gathering sufficient data which comprises training data. Training set's examples are put into a learning algorithm, which produces a trained "machine" that performs the specified task. The selection of a set of probable "machines," also known as the hypothesis class. A hypothesis class would be a neural network architecture (refer figure 2.12) with learnable weights. Machine learning models convert data into predictions or decisions by exploiting the information supplied during training. [Sim + 18]

Machine learning can be divided into 4 types:

Supervised Learning: Supervised learning maps features in training dataset to the target. Feature can be selected from domain knowledge, global features like permutation entropy, approximate entropy, stationarity, or fourier components can also be derived for understanding the time series.

The basic steps of supervised machine learning are:

- (1) Obtain a dataset with features and target, then separate it into training, validation, and test datasets.
- (2) Use training and validation datasets to inform a model of the relationship between features and target.
- (3) evaluate model using the test dataset.

Unsupervised: Unsupervised learning, in contrast to supervised learning, aims to detect patterns in a dataset and categorize individual instances within the dataset. These algorithms are unsupervised because the patterns in a dataset that may or may not exist are not informed by a target and must be determined by the algorithm. Clustering, association, and anomaly detection are some of the most common unsupervised learning tasks. Clustering, as the name implies, divides instances in a dataset into distinct clusters based on specific feature combinations.

Semi-supervised and reinforcement learning: Semi-supervised learning is "happy medium" between supervised and unsupervised learning, and it is especially useful for datasets with both labeled and unlabeled data (i.e., all features are present, but not all features have associated targets). This situation usually occurs when labeling becomes too time-consuming or too expensive. Semi-supervised learning is frequently used for medical images, in which a physician labels a small subset of images and then uses them to train a model. This model is then used to classify rest of the dataset's unlabeled images. Labeled dataset is then used to train a working model, which should outperform unsupervised models in theory. Finally, reinforcement learning is arguably the most accurate representation of the human learning experience because it learns through trial and error rather than data alone. [Cho + 20]

2.3.1 Neural Network

An artificial neural network (ANN) is a computational model inspired by the structure of neural networks in the brain. In simplified models, brain is made up of a large number of basic computing devices (neurons) that are linked together in a complex communication network, allowing the brain to perform highly complex computations. Artificial neural networks are formal computation constructs based on this computation paradigm. Neural networks were proposed as a method of learning in the mid-twentieth century. It generates an efficient learning paradigm and has recently been shown to achieve cutting-edge performance on a variety of learning tasks. The caveat is that training such hypothesis classes of neural network predictors is a computationally difficult problem. A popular method for training neural networks is based on *gradient descent* framework which employs *backpropagation algorithm* to compute gradients. [SV14]

Feedforward neural network: Concept of neural networks (NN) is that many neurons can be linked together by communication links to perform complex computations. Structure of a neural network is commonly described as a graph, with nodes representing neurons and each (directed) edge connecting the output of one neuron to the input of another. Input of a neuron is obtained by taking a weighted sum of the outputs of all neurons connected to it. Each single neuron is modeled as a simple scalar function σ , σ is also called activation function. The figure 2.12 shows the neural network. Some of the activation functions are:

- Sigmoid (x) = $\frac{1}{1+e^{-x}}$

- $\text{Tanh}(x) = \frac{2}{1+e^{-2x}} - 1$
- $\text{ReLU}(x) = \max(0, x)$

Activation function is a nonlinear transformation of the input signal (x) that determines whether or not a neuron should be activated. This is critical for NN because a network without an activation function is simply a linear regression model that cannot handle complex tasks. [Qi + 19]

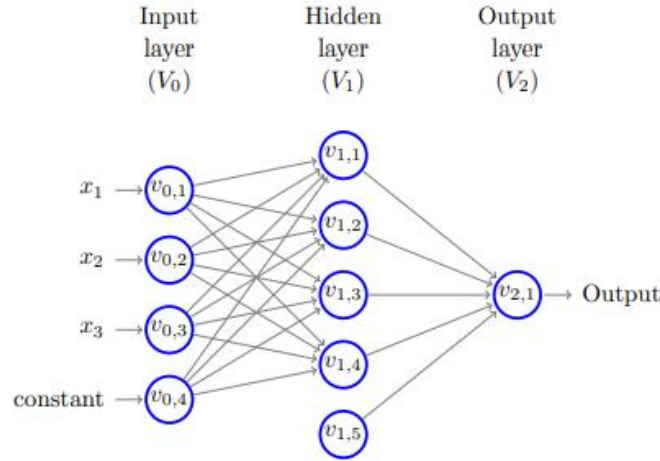


Figure 2.12: shows neural network [SV14].

Loss function: Model's parameters can be trained by calculating difference between the desired and actual outcomes. This measurement is known as the loss function, which is also known as objective or cost function. Loss function should be determined by the specific problem and frequently has a real-world interpretation. Root mean square error (RMSE) and the mean absolute error (MAE), for example, are both methods for calculating the distance between two vectors: the vector of predictions and the vector of target values. Formula for calculating RMSE and MAE is shown below.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - y_t)^2}{n}}$$

$$MAE = \frac{\sum_{i=1}^n |y_i - y_t|}{n}$$

where i is the sample index, y_i is the predicted value, and y_t is the targeted value. [Qi + 19; SV14]

Autoencoder reconstruction loss: Autoencoder is made up of two components: an encoder and a decoder. In a latent space Z , the encoder maps its input x to a representation z . During training, the decoder attempts to reconstruct x from z , ensuring that no useful information was lost during the encoding phase. In general, the reconstruction loss is defined as the distance $d_{AE}(x_i, f(x_i))$ between the autoencoder's input x_i and the corresponding reconstruction $f(x_i)$. One way to express it is to use the mean squared error of the two variables:

$$L = d_{AE}(x_i, f(x_i)) = \sum_i \|x_i - f(x_i)\|^2 \quad (2.1)$$

This loss function ensures that the learned representation retains important information from the initial one, allowing for the reconstruction. [Eli + 18]

The problem of overfitting and underfitting: An NN algorithm's main goal is to have good generalization ability, which is a measure of how accurately the algorithm can predict outputs from previously unknown data. However, overfitting or underfitting is a cause of poor NN algorithm performance. Overfitting occurs when the NN algorithm attempts to fit every data point in the training set, leaving the model vulnerable to noise or outliers. Underfitting, on the other hand, occurs when the NN algorithm fails to extract a reasonable relationship between the data points in the training set. Regularization and dropout are two techniques for avoiding overfitting and underfitting. [Qi + 19]

Gradient descent optimization algorithms: The training algorithm modifies the model's parameters to minimize the loss function; this learning is accomplished through gradient descent. Gradient descent can be used to train a model with many weights: at each iteration, the gradient of the loss function with respect to the weights is computed and the weights are updated. This translates into the update rule for one specific weight w .

$$\omega \rightarrow \omega' = \omega - \lambda \frac{\partial L}{\partial \omega_k} \quad (2.2)$$

where ω' is the updated weight, $\frac{\partial L}{\partial \omega_k}$ is the gradient of the loss function with respect to the weight ω , and λ is a parameter representing the length of the step to take in the gradient's direction, also known as the learning rate. The partial derivative of the loss with respect to each parameter is computed using the backpropagation algorithm, which employs the chain rules of derivatives to propagate the gradient to each parameter from the network's output to its input layer. To compute the gradient, the loss for each training example must be computed. Because doing this for each iteration of weight updates is computationally expensive, a common strategy is to estimate the gradient on a small sample of training examples chosen at random. This is known as stochastic gradient descent (SGD). A mini-batch is a collection of randomly selected inputs. [M16]

Challenges with vanilla mini-batch:

- Choosing an appropriate learning rate can be difficult. A low learning rate results in painfully slow convergence, whereas a high learning rate can obstruct convergence and cause the loss function to fluctuate around the minimum or even diverge.
- In addition, all parameter updates use the same learning rate. If the data is sparse and the features occur at different frequencies, we may not want to update them all to the same extent, but rather perform a larger update for rarely occurring features.
- Another significant challenge of minimizing highly non-convex error functions, is avoiding becoming trapped in their numerous suboptimal local minima. The difficulty is caused by saddle points, which are points where one dimension slopes up and the other slopes down. Because the gradient is close to zero in all dimensions, these saddle points are usually surrounded by a plateau of the same error, making it notoriously difficult for SGD to escape.

To overcome these challenges, momentum-based algorithms and flexible learning rate-based algorithms like Adagrad and Adam are used. [Rud + 16]

Evaluation metrics: A confusion matrix (table 2.2) displays the outcomes of a binary classifier versus the ground truth. Ground truth is the real target to be learned by a binary classifier.

		Ground Truth	
		Positive	Negative
Prediction	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

Table 2.2: Confusion matrix

There are only four possible results for a binary classifier: true positive (TP), positive prediction and positive ground truth; false positive (FP), positive prediction and negative ground truth; true negative (TN), negative prediction and negative ground truth; false negative (FN), negative prediction and positive ground truth. The most obvious indicator for classifier evaluation is accuracy, which corresponds to the number of properly categorized samples. Accuracy can be calculated using the results of the confusion matrix as:

$$Accuracy = \frac{\text{Correct Predictions}}{\text{Total Predictions}} = \frac{TP + TN}{TP + FP + TN + FN}$$

Because of the bias in the number of samples of positives and negatives in an imbalanced dataset, this metric of accuracy is useless. The true positive rate is defined as:

$$True\ positive\ rate = \frac{\text{Correct Predictions}}{\text{Total Predictions}}$$

2.3.2 Bayesian Optimization

The weights of the network can be learned using gradient descent, but deciding the architecture of the neural network like the number of neuron per layer to use, and the learning rate to use require lots of tuning of network architecture. Bayesian optimization (BO) is a powerful tool and versatile tool for hyperparameter tuning and more broadly for automatic design selection. This design process can be thought of as optimizing a black-box function that is difficult to evaluate for economic or computational reasons. Critically, the objective functions are unknown and costly to assess. The challenge is to find the maximum of such expensive objective functions in a few sequential queries while minimizing time and cost. The BO finds solution of function x^* :

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \quad (2.3)$$

by performing a series of x_1, \dots, x_T evaluations of f in order to find the optimum of f in the fewest iterations. It creates a surrogate for the objective and uses a Gaussian process to quantify the uncertainty in that surrogate.

Gaussian process: Bayesian optimization determines f from evaluations by constructing a Gaussian process. We can associate a normally distributed random variable at any point in the continuous input space thanks to this flexible distribution. We obtain the predictive distribution for a new observation x' , which has a Gaussian distribution as well.

Acquisition functions: Because the original function $f(x)$ is expensive to evaluate, surrogate model is used to construct a cheaper function, termed as acquisition function, to select the next point to evaluate. To select the next location, the acquisition function is maximized rather than the original function. [Ngu + 19]

$$\mathbf{x}_{t+1} = \arg \max_{\mathbf{x} \in \mathcal{X}} \alpha_t(\mathbf{x}) \quad (2.4)$$

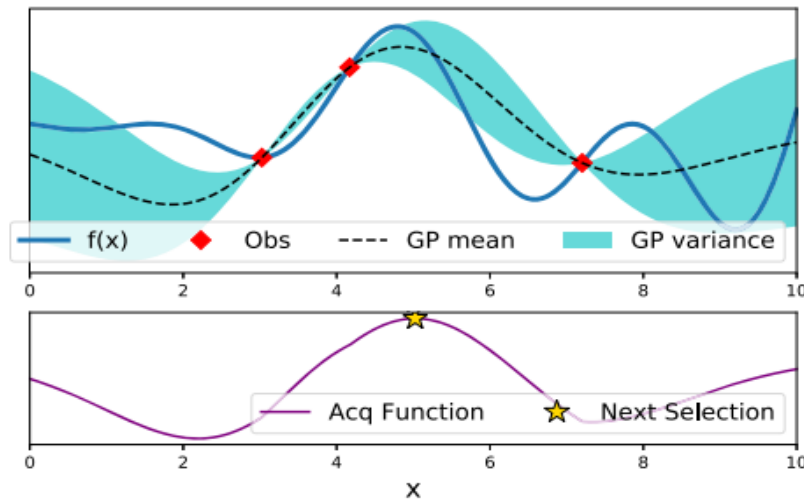


Figure 2.13: Shows example of gaussian process and bayesian optimization [Ngu + 19]

2.4 Anomaly Detection

An outlier is an observation (or group of observations) that appears to contradict the rest of the dataset, or "patterns in data that do not adhere to a well-defined idea of typical behavior." Many approaches for detecting anomalies have been proposed, including distance and density-based strategies as well as subspace or submanifold-based techniques. Because most of these techniques make no explicit use of time, they are less suitable for time series analysis. [Ahr + 19; Mal + 16] Deep learning i.e deep neural network learning methods are frequently being used for anomaly detection. A deep anomaly detection system outperforms shallow neural networks. [Cha + 19] Deep anomaly detection strategies can be defined based on label availability as:

(1) Supervised deep anomaly detection: This entails training a deep supervised binary or multiclass classifier using labeled data from both anomalous and non-anomalous incidences. Because of the scarcity of labeled training data, supervised approaches are not widely used. Also, supervised models are incapable of detecting an unseen example that does not belong in the training data. [Tan + 19]

(2) Semi-supervised deep anomaly detection: Semi-supervised approaches are more extensively employed since labels of normally occurring examples are significantly easier to get than labels of anomalies. These strategies use normally positive class to differentiate outliers. When training is done in a semi-supervised fashion i.e. on data with no anomalies, a deep autoencoder produces reduced reconstruction error for normal instances over abnormal instances.

(3) Unsupervised deep anomaly detection: Detection of unsupervised deep anomaly methods is solely based on the intrinsic properties of training data. These approaches are preferable because training data is typically imbalanced and anomalous data is far less common. These methods are frequently used for automatic labeling of unlabeled data

samples. Unsupervised approaches are further classified as clustering-based and reconstruction-based. [Cha + 19; Guo + 18]

Deep anomaly detection algorithms' outputs:

1. **Score:** Output of the deep anomaly detection model can be a score that specifies how likely the data point is anomalous. A threshold can be set to classify anomalies with scores above this threshold as anomalous and others as not anomalous, as in the case of unsupervised anomaly detection.
2. **Label:** In this situation, output of the anomaly detection system is a label indicating whether the data is normal or abnormal. [Cha + 19]

2.5 Dimensionality Reduction

Dimension reduction is a critical step in analyzing enormous high-dimensional datasets effectively. It could be the primary goal of the analysis to visualize high-dimensional data, or it could be an intermediate step that enables subsequent analyses, such as clustering. Principal component analysis (PCA) is likely the oldest and most widely used approach for calculating lower-dimensional representations of multivariate data. The technique is linear in the sense that the components are linear combinations of the original variables (features), but the data's non-linearity is kept for effective visualization. The method is described as an iterative computation of the direction of greatest variation followed by projection onto the perpendicular hyperplane. This soon yields a few perpendicular directions that account for most of the variation in the data, resulting in a low-dimensional representation of the data.

2.6 3D Coordinate Reference Changes

If reference frames for the CT scan of the printed WAAM part and WAAM robot path are not defined at the same points then both of the reference frames need to be aligned. In this work, the different reference frames were aligned using a homogeneous transformation matrix. To translate a vector from Frame 1 to Frame 0, a homogeneous transformation matrix A_1^0 can be utilized:

$$A_1^0 = \begin{bmatrix} R_1^0 & o_1^0 \\ \mathbf{0}^T & 1 \end{bmatrix}$$

R_1^0 be the rotation matrix of Frame 1 with respect to Frame 0. Let o_1^0 be the vector describing the origin of Frame 1 with respect to Frame 0. The transformation of a vector from Frame 1 to Frame 0 is easily expressed by a single matrix that contains the rotation matrix of Frame 1 with respect to Frame 0 and the translation vector from the origin of Frame 0 to the origin of Frame 1. As a result, the coordinate transformation can be rewritten compactly as:

$$\tilde{p}^0 = A_1^0 \tilde{p}^1 \quad (2.5)$$

Homogenous transformation matrix A_1^0 can be used to describe transformation between Frame 0 and Frame1. The transformation matrix satisfies the equation below:

$$\tilde{\mathbf{p}}^1 = \mathbf{A}_0^1 \tilde{\mathbf{p}}^0 = (\mathbf{A}_1^0)^{-1} \tilde{\mathbf{p}}^0 \quad (2.6)$$

Rotation matrix:

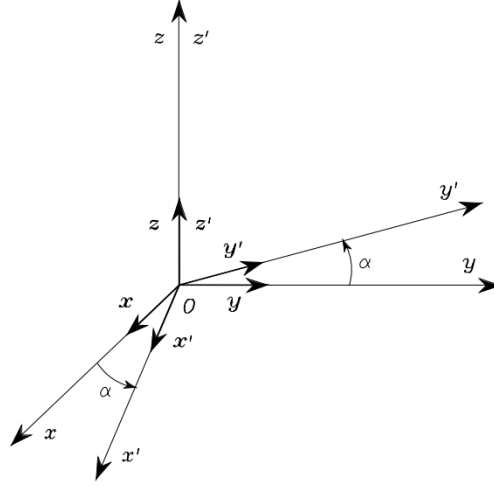


Figure 2.14: Rotation of 0-xyz by an angle α about axis z [BLL10].

Rotation matrix of frame 0- $x'y'z'$ with respect to frame 0-xyz is [BLL10]:

$$\mathbf{R}_z(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2.7 Algorithms

In the following sections, the main algorithms used for developing anomaly detection models are introduced.

2.7.1 Convolutional Neural Network

Convolutional neural networks (CNN) are Inspired by biology, specifically the structure of the animal visual cortex. They are useful for applications in regular-grid data such as images, and time series, if locality and shift-equivariance/invariance of feature extraction is desired [Eli + 18]. Convolutional neural networks are based on three fundamental concepts: local receptive fields, shared weights, and pooling. Each of the concepts is defined below. [M16]

- **Local receptive files:** Hidden neuron analyzes its specific local receptive field. We can move the local receptive field to the right by adjusting the stride parameter. The figure 2.15 shows the local receptive field for the first hidden neuron.
- **Weights and biases shared:** Each hidden neuron has biases and weights associated with its local receptive fields. A kernel or filter is often said to be defined by its shared weights and bias.
- **Pooling layers:** Pooling layers are used to reduce the dimensions of the hidden layer by combining the outputs of neuron clusters at the previous layer into a single neuron in the next layer

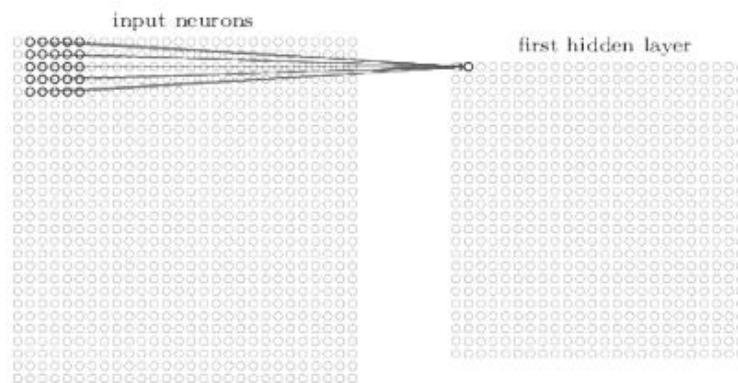


Figure 2.15: Showing local receptive field of the first hidden neuron. [M16]

Sharing weights and biases has the significant advantage of reducing the number of parameters involved in a convolutional network. Because of its parameter efficiency, CNN can be a good alternative for both univariate and multivariate time series data. DeepAnt uses unsupervised methods to detect anomalies in time series data using CNNs, the DeepAnt architecture is shown in the figure 2.16. CNN is also more resistant to variation than other neural networks and statistical models. [Mun + 19]

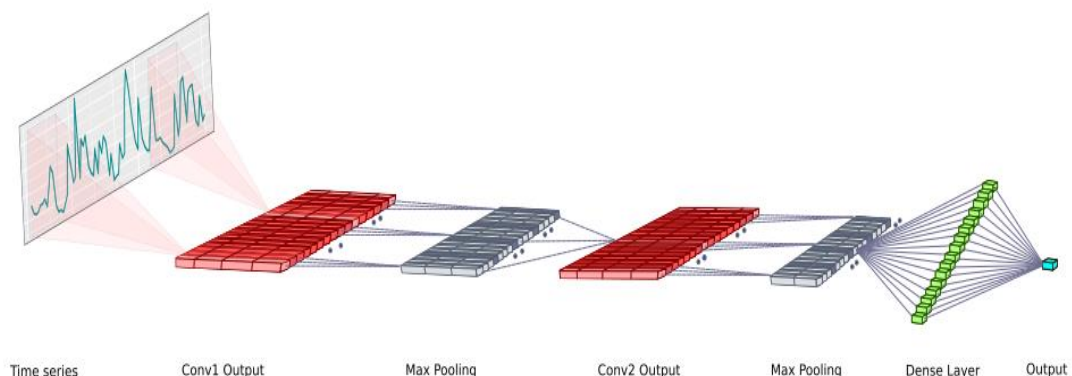


Figure 2.16: Showing DeepAnt architecture proposed by Munir et al. [Mun + 19]

Tan et al. [Tan + 19] developed an encoder-decoder architecture for identifying anomalies in sequence image data collected during the Laser additive manufacturing pilot system (LAMPS) testbed, a platform that leverages selective laser sintering (SLS) technology for AM. They designed a CNN-based encoder network, shown in the figure 2.17 and trained it in unsupervised fashion using nearly defect-free input.

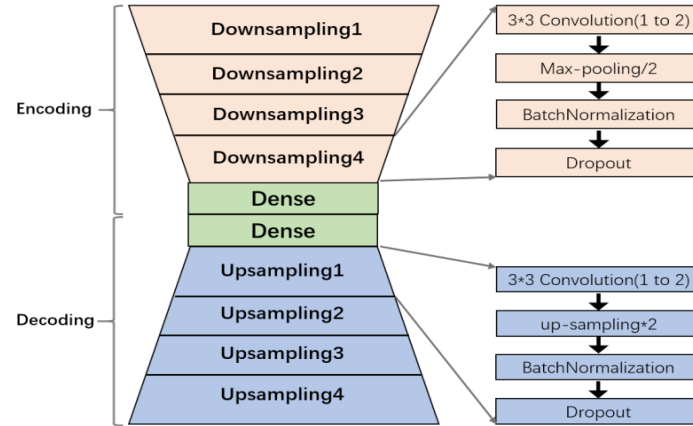


Fig. 3: Network Structure

Figure 2.17: Showing convolutional autoencoder architecture developed by Tan et al. for anomaly detection. [Tan + 19].

2.7.2 Autoencoder

Autoencoders were first introduced as a type of neural network that can be trained to reconstruct its input. Their primary goal is to learn an "informative" representation of the data in an unsupervised manner that can be used for various implications such as clustering. The problem can be formally defined as:

$$\arg \min_{A,B} E[\Delta(\mathbf{x}, B \circ A(\mathbf{x}))] \quad (2.7)$$

Where A is encoder, B is decoder and E denotes the expectation over the distribution of x and is the reconstruction loss function, which measures the distance between the decoder's output and the input. The autoencoder model is illustrated in the figure 2.18.

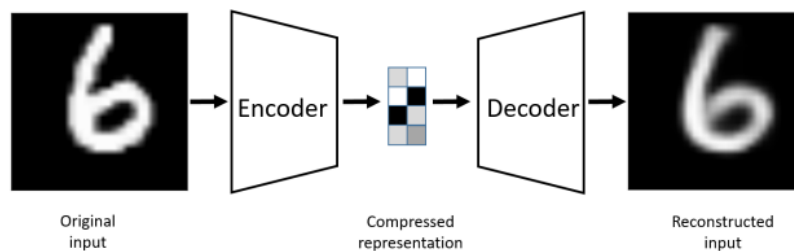


Figure 2.18: Showing example of autoencoder. [Ban + 20]

A and B are neural networks in the most common form of autoencoders. A linear autoencoder is obtained when A and B are linear operations. In the case of a linear autoencoder, removing the non-linear operations results in the same latent representation as Principal Component Analysis (PCA). As a result, an autoencoder is a generalization of PCA in that, rather than finding a low-dimensional hyperplane in which the data lies, it can learn a non-linear manifold. The main approach in anomaly detection tasks involving AEs is to train on normal, non-anomalous data instances, so that reconstruction of anomalous data results in a high reconstruction error. The reconstruction error is minimized during the training phase. Anomaly data points are used in the testing phase to validate the anomaly detection capabilities, and if they exceed a predefined or dynamically calculated threshold, they are labeled as anomalies. [Ban + 20]

2.7.3 Inception network

The Inception network shown in the figure 2.20 is a model used for time series prediction and classification, constructed by cascading multiple Inception modules as seen in the figure 2.19. An Inception module's central notion is to apply many filters to an input time series at the same time. The module comprises filters of varying lengths, which allows the network to automatically extract useful features from both long and short-time series.

An Inception network classifier is made up of two distinct residual blocks. Each block is made up of three Inception modules rather than standard fully convolutional layers. The input of each residual block is sent via a shortcut linear connection to be added to the input of the next block, so minimizing the vanishing gradient problem by permitting a direct flow of the gradient. Following these residual blocks is a global average pooling (GAP). Finally, a typical fully connected layer with a number of neurons equal to the number of classes in the dataset is used. Python library tsai [Tsa + 2022] has InceptionTime Plus, which has the capacity to forecast time series. [Ism + 20]

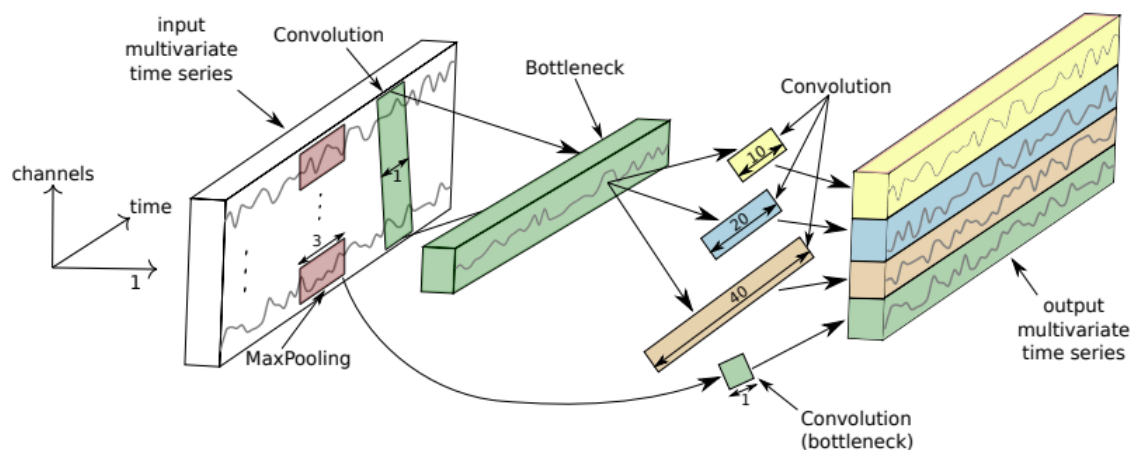


Figure 2.19: Showing inception module in Inception network [Ism + 20]

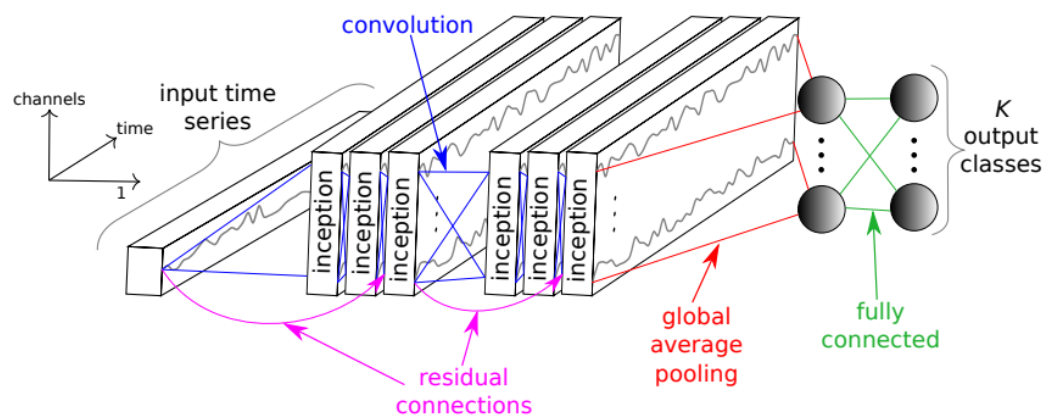


Figure 2.20: Showing Inception network with its components [Is_m + 20]

3 Experimental Setup, Data Creation & Processing

In this chapter lab and factory sensor and setup, different kinds of data handled and analyzed, the experiment conducted, training data for anomaly detection model, and some initial data analysis are discussed.

3.1 Lab & Factory setup

WAAM setups typically include four components: the welding source, the kinematic system (robot in our case), the control system, and the monitoring system. A Robot-based WAAM setup can be seen in figure 3.1. [Rei + 20b] In this work, experiments were performed with steel alloy as feedstock material. The different systems and sensors involved in Siemens WAAM lab and factory are stated in the following section.

3.1.1 Lab Setup & Sensors

Lab setup and sensors are stated in this section.

Sensors:

Sensor	Sample rate	Output
Current	4 kHz	0.0 – 10.0 amps
Voltage	4 kHz	0.0 – 10.0 volts
Gas Flow	20 kHz	Scaler value in liter/min

Table 3.1: Sensor details (lab).

Process setup:

- **Welding Source:** Fronius CMT 4000 advanced
- **Wire feeder:** VR 7000

Edge device: Sensors are connected to edge devices which take care of data acquisition, processing, and visualization.

- **Model:** Nvidia Jetson AGX Xavier,
- **CPU:** 1 x ARMv8.2 @ 2.26 GHz

3.1.2 Factory Setup & Sensors

Factory setup and sensors are stated in this section.

Sensors:

Sensor	Sample rate	Output
Current	1 kHz	0.0 – 200.0 amps
Voltage	1 kHz	0.0 – 40.0 volts
Gas Flow	20 Hz	0.0 – 30.0 liter/min

Table 3.2 Sensor details (factory).

Process setup:

- **Welding Source:** Fronius TPS/i 4000 advanced
- **Wire feeder:** VR 7000

Edge device:

- **Model:** Nvidia Jetson AGX Xavier,
- **CPU:** 1 x ARMv8.2 @ 2.26 GHz

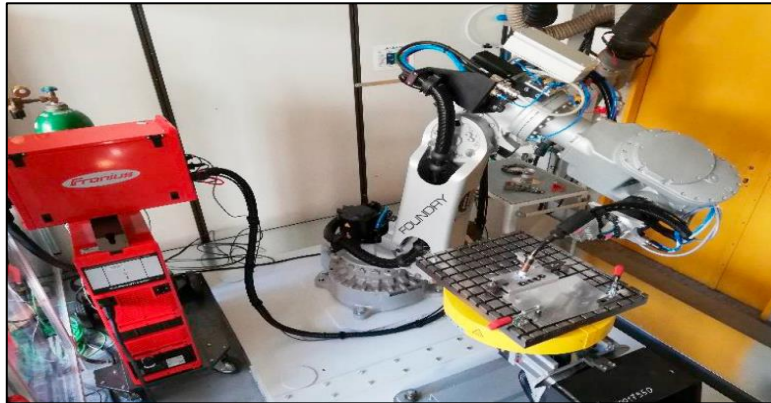


Figure 3.1 : WAAM system cell (Siemens lab). [Rei + 20b]

3.2 Different Kinds of Data

In this work, 4 different WAAM printed structures and their voltage data were analyzed. Material used for deposition were steel alloys. 3 structures were printed in Siemens lab, Munich, and 1 was printed in Siemens factory, Görlitz. Different structures were printed with different process parameters. Overview of the different structures is shown in table 3.3, in the following paragraph, different structures are discussed in detail.

Wall no.	Structure type	Intentional process parameter changes	Acquired from	Layer structure
1	Linear wall	Reduced gas flow and variation in robot head speed	Siemens lab, Munich	Linear
2	Oxidized complex wall	Reduced gas flow in some layers	Siemens lab, Munich	Meander and Contour
3	Inverted pyramid complex wall	None	Siemens lab, Munich	Meander and Contour
4	Airfoil blades	None	Siemens factory, Görlitz	Airfoil blades

Table 3.3: Different kinds of data.

3.2.1 Linear Wall (wall no. 1)

The linear wall shown in figure 3.2 was built using linear layer deposition strategy, voltage waveform (shown in figure 3.3) of linear wall is different from other structures analyzed in this work, the waveform does not have phases as described in the CMT (section 2.1.1). In this wall, gas flow was reduced, and robot head speed was varied to understand the variation in voltage waveform when the process parameters were changed.

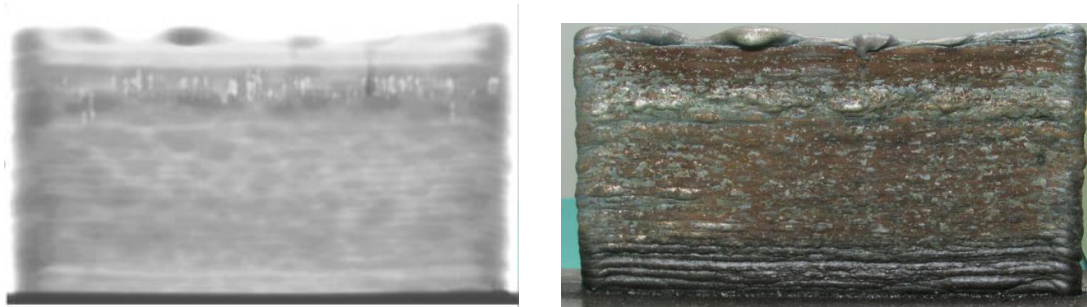


Figure 3.2: Linear wall and its x-ray (left), printed in the lab.

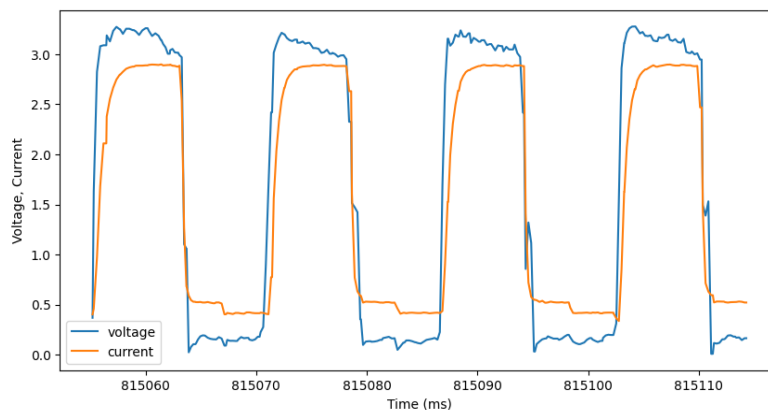


Figure 3.3: Showing voltage and current waveform.

3.2.2 Oxidized Complex Wall (wall no. 2)

One layer of this wall is a combination of 2 meanders (figure 3.4) and 1 contour (boundary of meander). In this wall, gas flow was reduced partly in layer 11 and 15 to produce oxidized layers. Voltage waveform in this wall is different from linear wall, voltage waveform for this wall follows the CMT cycle as mentioned in section 2.1.1. and is shown in figure 3.6. Oxidized complex wall and its CT scan can be seen in figure.3.5.

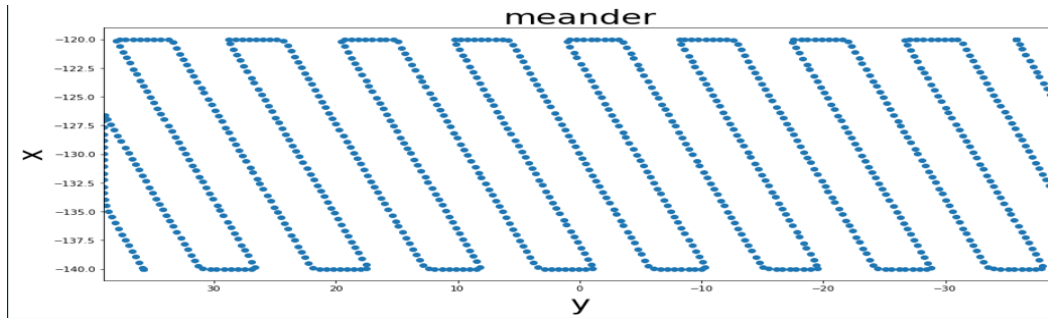


Figure 3.4: Robot path during meander.

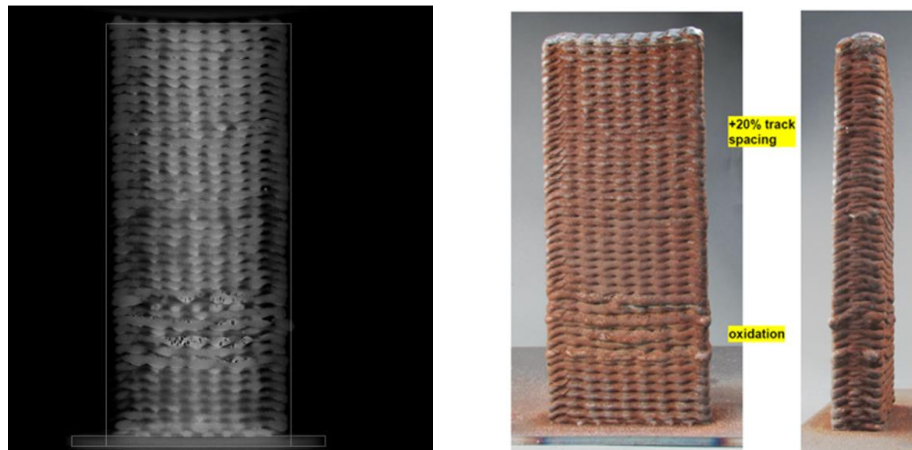


Figure 3.5:CT scan (cross-sectional) and part view (wall no. 2).

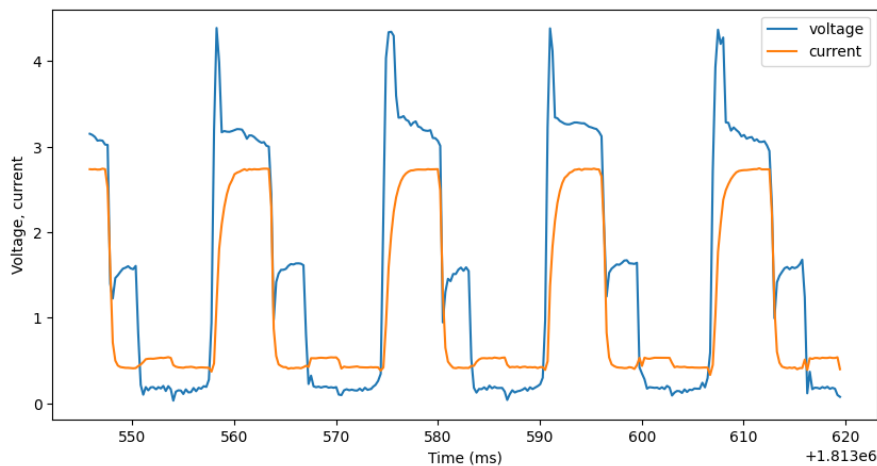


Figure 3.6: Showing voltage and current waveform (wall no. 2).

3.2.3 Inverted Pyramid Complex Wall (wall no. 3)

This wall was built using the same combination of layers as an oxidized complex wall and had same current and voltage cycle as oxidized complex wall (wall no. 2). In this wall, an overhanging structure was printed. Even though no intentional process parameters were changed to create defects the wall has large cavities and pores as can be seen in the figure 3.7. These defects are undesirable.

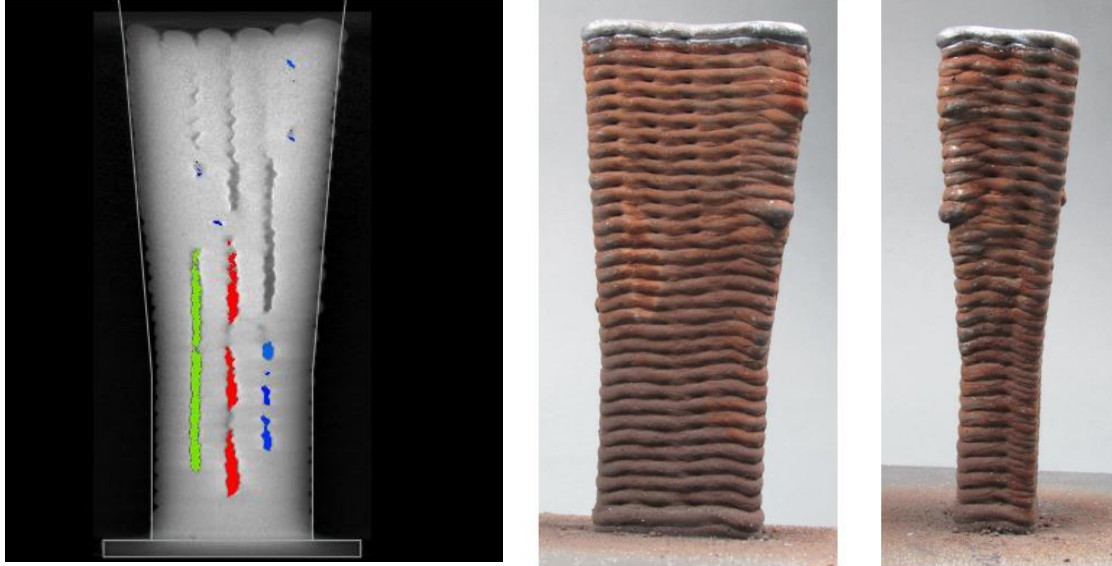


Figure 3.7: CT scan (cross-sectional) and part view (wall no. 3).

3.2.4 Airfoil Blades (wall no. 4)

This structure was built in Siemens factory. One layer of the structure can be seen in figure 3.8. Voltage and current cycles can be seen in figure 3.9. Voltage and current cycle are different from all other walls printed in the lab due to the different WAAM printing setup utilized in production.

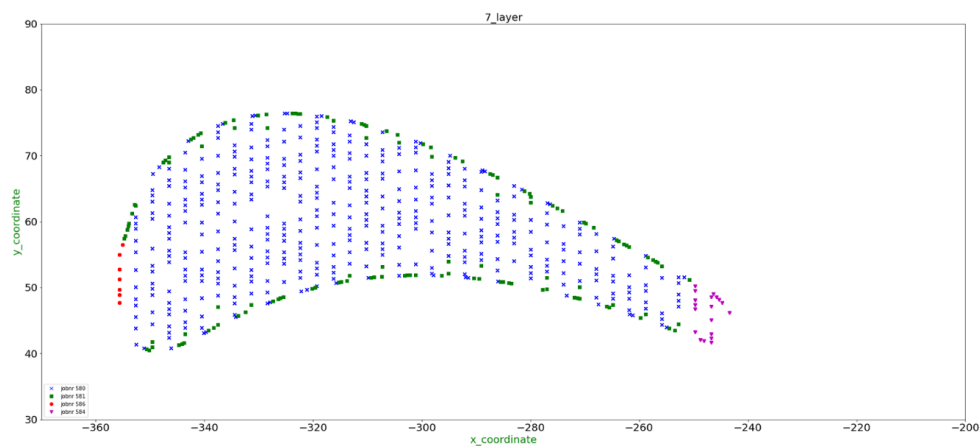


Figure 3.8: Showing robot path for printing airfoil for a particular layer.

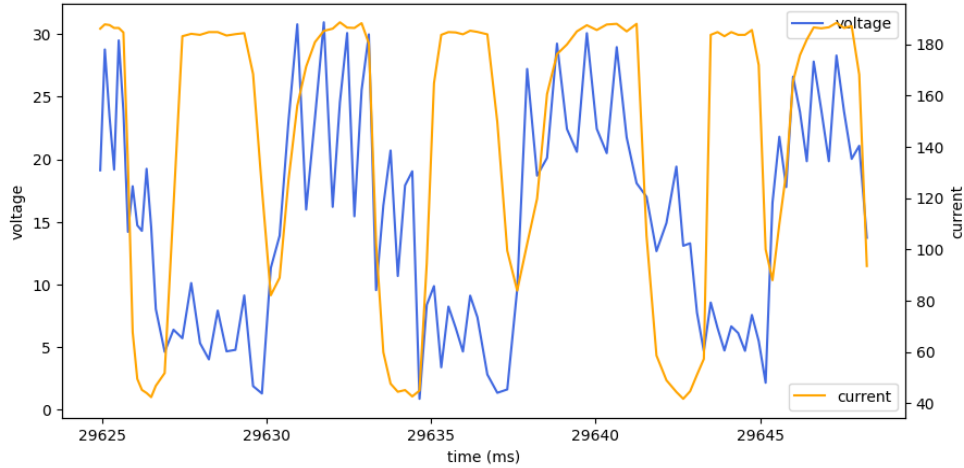


Figure 3.9: Showing voltage and current waveform (wall no. 4).

Primary focus of my work is on complex walls i.e., wall no. 2 and wall no. 3, which were printed in the lab, some initial analysis of linear wall and airfoil structure is also done in the thesis.

3.3 Data Processing And Labeling

In this section spatial labeling approach and initial data analysis are explained.

3.3.1 Data Cleaning and Preprocessing

Raw data collected from the sensors had many missing values and a lot of data was incorrectly recorded or not at all recorded. A thorough inspection of recorded data for each wall was carried out to clean and preprocess the data to make it useful for further analysis.

3.3.2 Labeling

Details of each defect like area, projection in z-direction, area in x-y plane, and center of the defect along with a CT scan of the printed structures were available for analysis. CT scan is necessary because the 3D wall structure cannot be inspected for defects manually with the human eye. CT scan was analyzed using the software MyVgl [myv], a screenshot of the CT scan in MyVgl is displayed in figure 3.10.

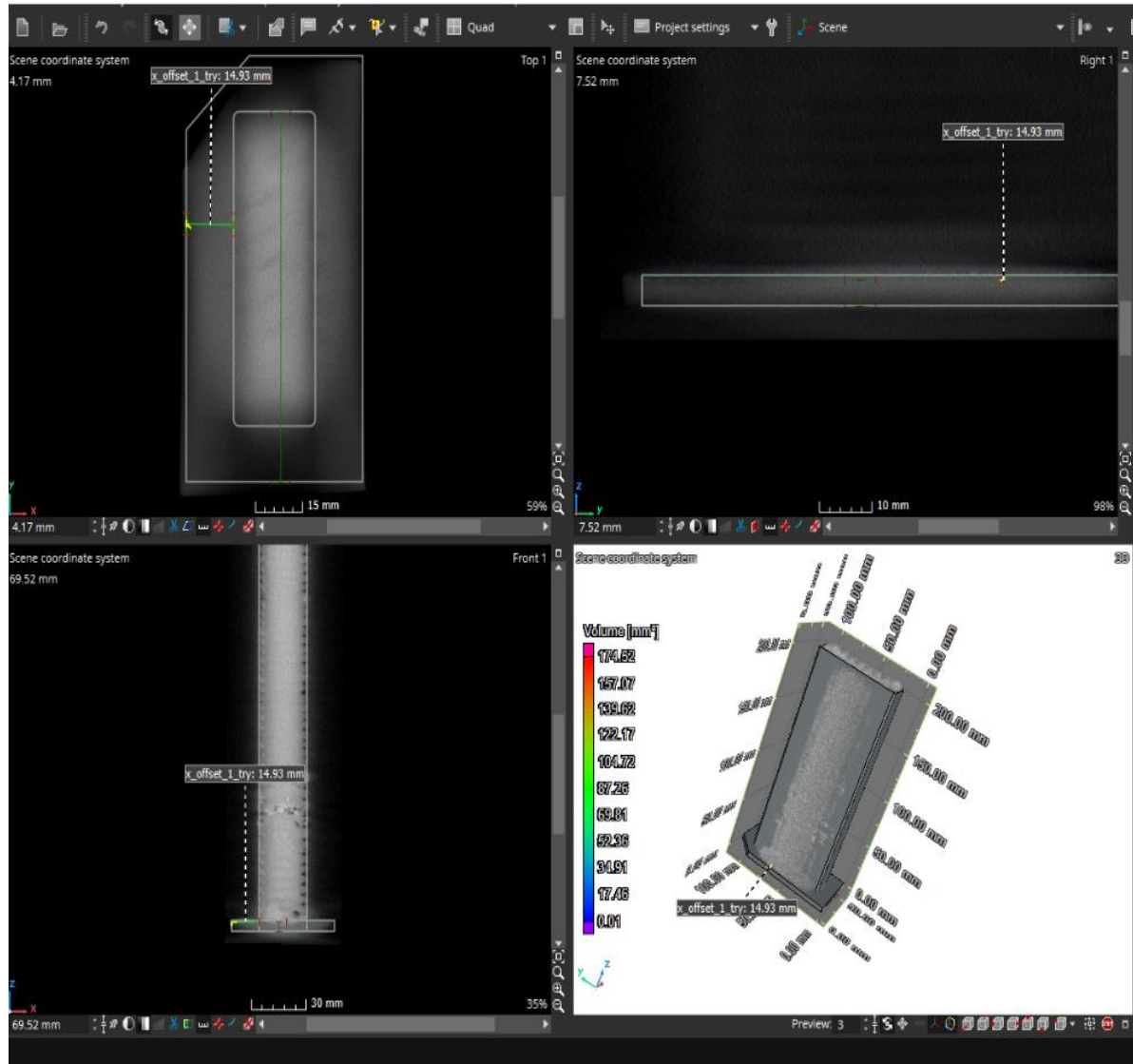


Figure 3.10: CT scan analysis. [myv]

To align the robot reference frames with the CT scan reference frames, rotation, and translation on both reference frames were applied using homogenous transformation matrix (mentioned in section 2.6). The parameters for both translation and rotation were manually selected by inspecting both robot and CT scan reference frames.

Visual inspection of the CT scan was used to determine height of each layer of the meander. Because meander layers are covered by contours from sides and printed layers are not of uniform height throughout, height of layers are imprecise, to account for this uncertainty margins in the z-direction are included in evaluation method discussed later in section 4.1. The figures 3.11 and 3.12 shows cross sectional view of meander layers taken from CT scan overlayed with defect locations and robot's path after reference changes.

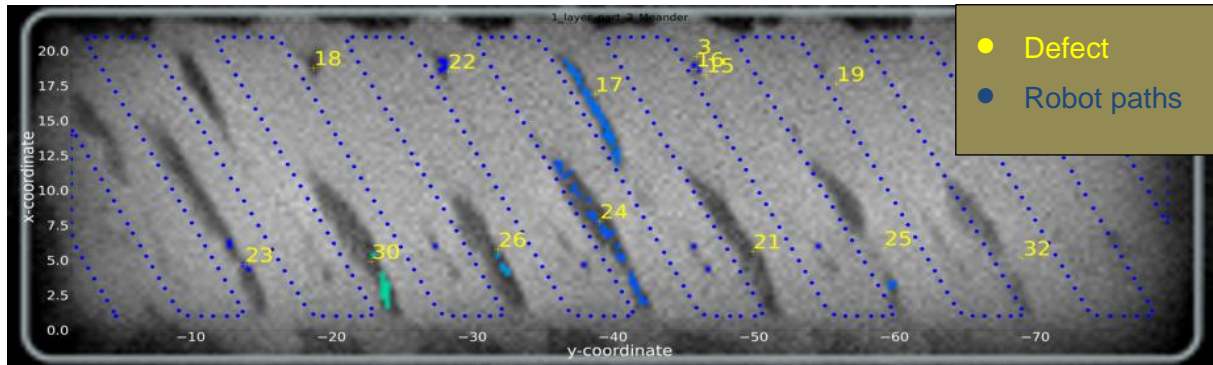


Figure 3.11: XY cross-sectional view wall no. 2 (CT scan).

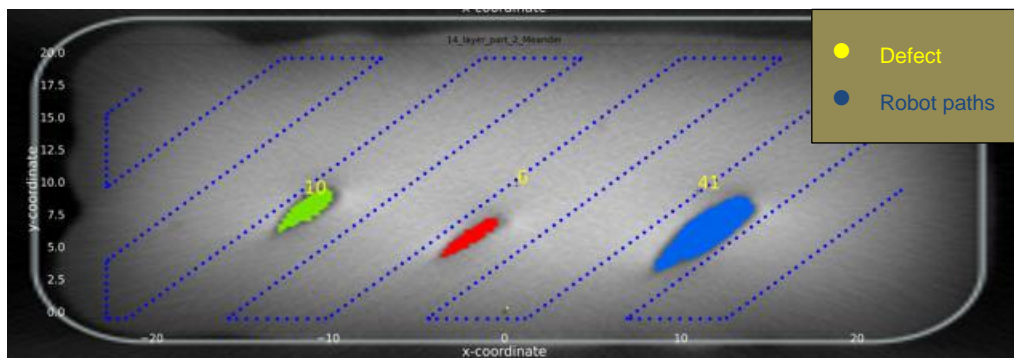


Figure 3.12: XY cross-sectional view wall no. 3 (CT scan).

3.4 Initial Data Analysis

Linear walls were first analyzed to learn more about the data. Global features like entropy and standard deviation were computed for the linear wall data using the *tsfresh library* [Chr + 18] to get an initial understanding of the data and its relation to gas flow changes. Features were estimated for small windows of raw voltage data. Box plots in figure 3.13 demonstrate the standard deviation for small windows of voltage data in various layers. Standard deviations of the short window of voltage data in low gas flow levels is larger than in other layers.

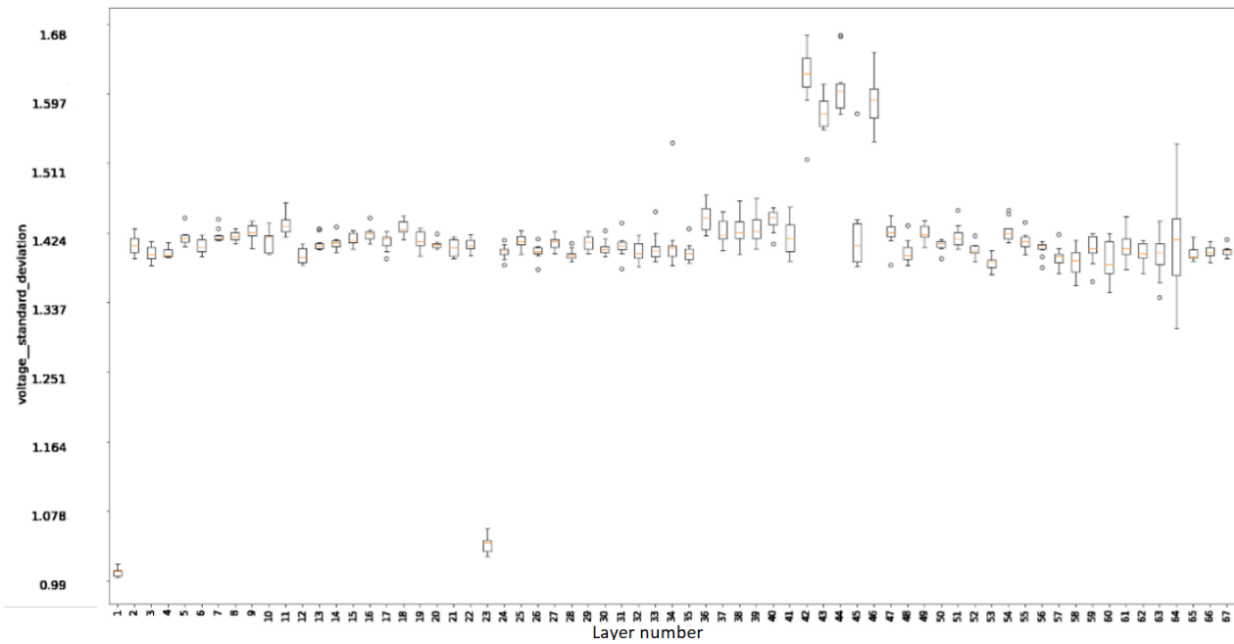


Figure 3.13: Boxplots of standard deviations for all layer.

3.5 Training Data Creation

Through visual inspection of the CT scan of wall no. 2, a few layers were found to have very few defects particularly layer number 19 and 20. Figure 3.14, shows a cross-sectional view of the CT scan of layer number 19.

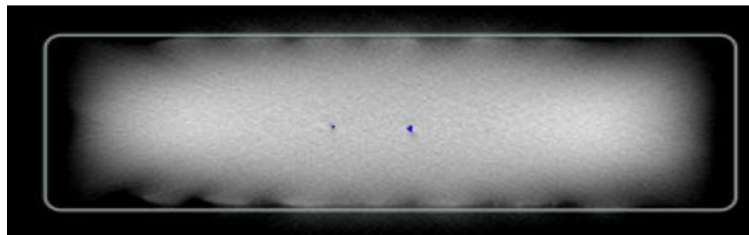


Figure 3.14: A cross-sectional view of layer number 19.

Unsupervised learning was utilized to learn non-anomalous cluster of data from the voltage data in layers 19 and 20. Since there are still some contaminations in layers the amount of contamination in voltage data was manually set while conducting experiments with anomaly detection models described in the following chapter.

4 Methodology for Lab Data

This chapter discusses the methodology used to develop anomaly detection models and evaluation algorithm used to measure the performance of developed models. Implementation in this chapter is done for lab data in particular wall no. 2 and wall no. 3.

4.1 Model Evaluation Algorithm

Because the commonly used evaluation metrics in classification cannot be used to evaluate the performance of models due to a lack of labeled data, an evaluation algorithm was developed that essentially checked the 3D spatial correlation between the locations of model anomalies, i.e., anomalies detected by the model in voltage data and the real defect, implying that if the model anomaly was found to be near the defect, it is considered true positive and otherwise false positive.

Because shape of the spread of part's defect through information of its area was unclear, they were considered cylindrical shaped for the evaluation, with the radius and height of the cylinder found by already available data of defect's area and projection in the z-direction.

Algorithm first determines whether any defect is within 2.5 mm in both positive and negative z-direction of the height of each model anomaly, i.e., height of the layer containing the model anomaly, determined by manual CT scan inspection. This total margin of 5 mm aids in dealing with uncertainty in the height of each layer in the z-direction. After this, it checks if the 2D Euclidean distance between the center of the defect (x, y coordinate) and the location of the model anomaly (x, y coordinate) is less than $4\text{mm} + \text{radius of a defect}$. Margin of 4 mm was used to deal with uncertainty in x-y coordinates caused by manual reference changes and varying centers for large continuous defects in each layer because the available data for the center of each defect is global center of the entire defect but not for the center per layer. Pseudo code of the algorithm is shown below:

Algorithm for model evaluation	
Input:	
1.	A: Data table with model anomalies along with x, y and z coordinates per instance of data
2.	B: Data table with center, lower and upper height, and radius of defects
Output:	
1.	Updated A with information about true positive or false positive per instance of model anomalies
2.	Updated B with information on whether the defect found a correlation with a model anomaly or not.
For each element i, in A do:	
For each element j, in B do:	
If j is in between 2.5 mm below and 2.5 mm above i's center, then	
If Euclidean distance between center of i and j is less than radius of j plus 4 mm, then	
Mark i as true positive and j as being detected	
else	
Mark i as false positive and j as not detected	
else	
Mark i as false positive and j as not detected	
end	

4.2 Feature Engineering

The voltage CMT cycle of both wall no. 2 and wall no. 3 data have different phases (refer section 2.1.1), these phases vary a lot in normal printing and anomalous printing. Figure 4.1 shows raw voltage and current CMT cycles in regions with low shielding gas flow and normal process parameters in wall no. 2., voltage in the anomalous region has higher voltage values in the peak phase compared to normal process.

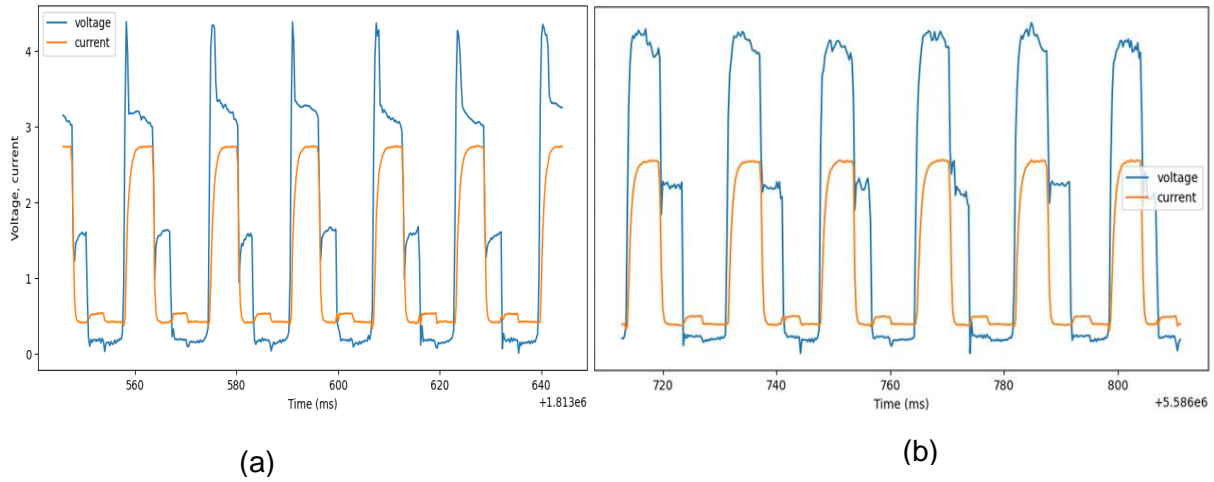


Figure 4.1: Instance of current and voltage CMT cycles during a) normal b) low gas flow region (oxidized).

To better capture information about voltage data, an algorithm was developed to compute features based on different phases of CMT cycle. Analysis of phase transition, start and end of both current and voltage CMT cycles was done for this. This manual selection yielded good segmentation of voltage data for all cycles except some very abnormal cycles which do not follow thresholds for phase transition. Figure 4.2 shows result of segmentation by the algorithm.

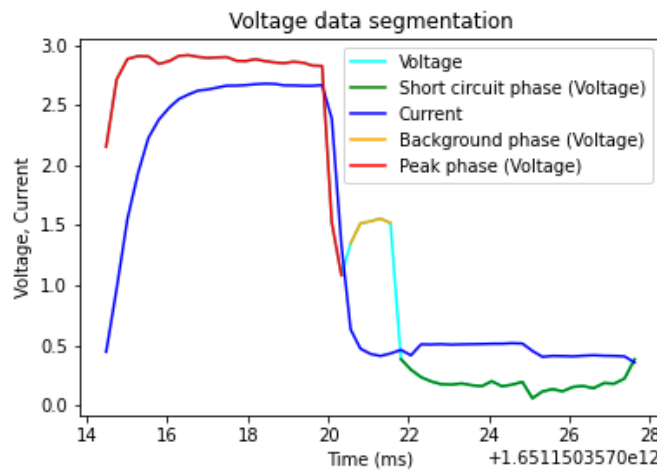


Figure 4.2: Voltage CMT cycle's phases segmentation using developed phase transition thresholds based algorithm.

Using same algorithm to segment the raw signal of current and voltage, different features were computed for voltage data. Current signal is mostly stable throughout the welding process irrespective of normal or abnormal printing because it is controlled by the welding equipment and therefore not analyzed. Table 4.1 below lists the features and their description.

No.	Feature name	Description
1	Max_whole_cycle	Max value in whole voltage cycle
2	Mean_whole_cycle	Mean value in whole voltage cycle
3	Mean_peak_phase	Mean value in peak phase of voltage cycle
4	Variation_peak_phase	Variance in peak phase of voltage cycle
5	Variation_background_phase	Variance in background phase of voltage cycle
6	Mean_background_phase	Mean value in background phase of voltage cycle
7	Len_cmt	Number of data points in whole voltage cycle
8	Len_background_phase	Number of data points in background phase of voltage cycle
9	Len_peak_phase	Number of data points in peak phase of voltage cycle
10	Duration_short_circuit_phase	Duration of short circuit phase of voltage cycle
11	Duration_background_phase	Duration of background phase of voltage cycle
12	Duration_cmt_phase	Duration of whole voltage cycle
13	Duration_peak_phase	Duration of peak phase of voltage cycle
14	Squared_sum_background_phase	Squared sum of voltage value in background phase of voltage cycle
15	Squared_sum_peak_phase	Squared sum of voltage value in peak phase of voltage cycle
16	Variation_whole_cycle	Variance in voltage data of whole voltage cycle.

Table 4.1: Voltage CMT cycle features and description.

4.3 Model Development Using Supervised Learning

Because of the uncertainty in defect origin locations, labeled data with normal and abnormal CMT cycles, particularly abnormal behavior in the case of pores and cavities, could not be obtained. An attempt was made to mark all oxidized data from wall no. 2, layer number 11 and 15 as defective and training data (described in section 3.5) as normal, and then use this labeled data set to develop a supervised learning-based model. Neural network and logistic regression were used as models, but model gave too many random anomalous cycles spreading throughout, even in layers with less anomalous data suggesting that the model trained with oxidized data as anomalous data is not reliable to detect defects other than oxidation. Therefore, unsupervised algorithms were attempted which will be discussed in the following sections.

4.4 Model Development Based on Forecasting Using Unsupervised Learning

A forecasting model to forecast the voltage signal was developed, which uses the mean squared error (forecasting error) of the forecasted signal and original signal to classify the chunk of the signal as anomalous or non-anomalous, threshold for anomaly classification was selected by setting the amount of contamination in the training data, and then taking lowest forecasting error for the contaminated data. Final contamination amount was decided based on the model's performance on that particular threshold.

Training data was taken as described in section 3.5. For this approach, python library *tsai* [Tsa + 2022] was used, in particular, InceptionTime model (discussed in section 2.7.3) was used because of its time series forecasting ability. 200 steps of voltage data or nearly 4 voltage CMT cycles were used as input and the next 20 steps or nearly the next half voltage CMT cycle was the forecasted signal, figure 4.3 shows original signal to be forecasted and forecasted signal by the developed model.

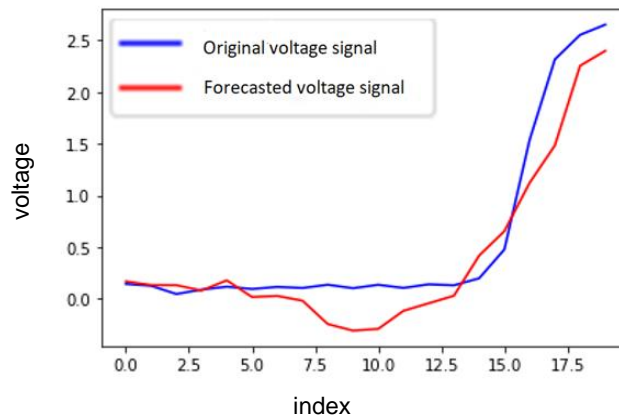


Figure 4.3: Showing forecasted and original signal prediction by trained Inception network.

4.5 Model Development Based on Single CMT Cycle Using Unsupervised Learning

In this approach, models were developed to analyze each voltage CMT cycle. Models were trained in an unsupervised manner. Training data as described in section 3.5 was used.

Two different approaches were tried here:

1. **Reconstruction-based:** In which the raw voltage CMT cycle was learned by the model.
2. **Feature-based:** In which the features of single voltage CMT cycles were learned by the model.

In the following sections, these approaches are discussed.

4.5.1 Reconstruction-based learning

Since convolutional networks have been found to have good prediction, and feature extraction capability in the time series. A convolutional network with a flattening layer and dense layers in end was developed and trained. Network was trained to learn the raw normal voltage CMT cycles and reconstruct them. Anomalous cycles will give higher reconstruction loss compared to normal cycles. Threshold for anomaly detection was set by adjusting the amount of contamination as described in section 4.4.

Hyper tuning to find the best architecture of the network was done using Bayesian optimizing techniques (refer section 2.3.2). Raw signal was padded so that the input was of constant size even for varying length voltage signals for a single CMT cycle. Length of input was set to 140. Input and reconstructed signal by the convolutional network is shown in the figure 4.5. The model architecture is shown in figure 4.4.

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv1d_1 (Conv1D)	(None, 111, 30)	930
flatten_1 (Flatten)	(None, 3330)	0
dense_1 (Dense)	(None, 140)	466340

Total params: 467,270
 Trainable params: 467,270
 Non-trainable params: 0

Figure 4.4: Architecture of reconstruction-based convolutional network.

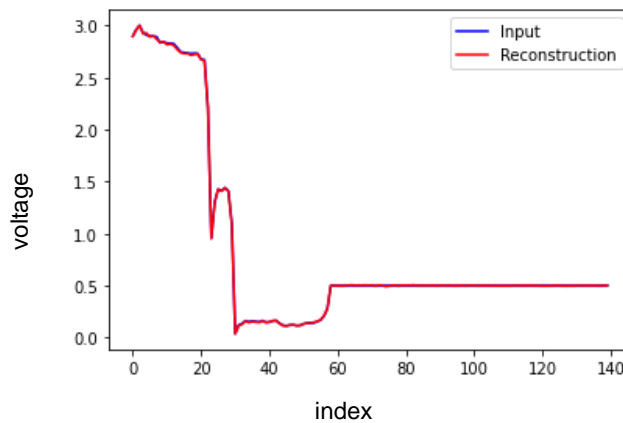


Figure 4.5: Instance of original input and reconstruction by the developed convolutional network.

4.5.2 Feature-based learning

Since autoencoder can compress and reconstruct the input features, an autoencoder was trained. In this approach instead of the raw voltage CMT cycle, network was trained to learn features of the CMT cycle and reconstruct them. Any CMT cycle with very different feature values from normal training data will have higher reconstruction loss. Threshold for anomaly detection was set by adjusting the amount of contamination as described in section 4.4. By checking the correlation between input features only 12 independent variables were selected as input for this model.

Hyper tuning to find the best architecture of the autoencoder network was done using Bayesian optimizing techniques (refer section 2.3.2). Architecture of the model is shown below (figure 4.6), dropout layers were also included in the network to reduce overfitting while training the network.

Model: "sequential_1"		
Layer (type)	Output Shape	Param #
dense_4 (Dense)	(None, 11)	143
dropout_2 (Dropout)	(None, 11)	0
dense_5 (Dense)	(None, 10)	120
dense_6 (Dense)	(None, 10)	110
dropout_3 (Dropout)	(None, 10)	0
dense_7 (Dense)	(None, 12)	132
Total params: 505		
Trainable params: 505		
Non-trainable params: 0		

Figure 4.6: Architecture of feature based autoencoder.

4.6 Model Development Based on Multiple CMT Cycles Using Unsupervised Learning

To obtain more temporal and spatial information, an attempt was made to combine features for 6 consecutive cycles by averaging and considering the variance for feature values for each feature discussed in section 4.2; this will provide information about the average feature value for 6 consecutive cycles and their variation. An autoencoder was used to learn features with input and output consisting of 32 features i.e., average and variance of the features of 6 consecutive cycles. Shin et al. [Shi + 20] also used a similar approach to detect pores, this approach lets analysis of multiple CMT cycles together. Model was trained with data as described in section 3.5 and anomaly threshold was selected in the same way as described in section 4.4. The average value of coordinate-x and coordinate-y for consecutive 6 cycles was

taken as the location of chunk of the data. Hyper tuning to find the best architecture of the autoencoder network was done using Bayesian optimizing techniques. Some of the features showed a clear correlation with regions where gas flow was reduced i.e., layers 11 and 15 in wall 2, as shown in the figures 4.7, 4.8 and 4.9.

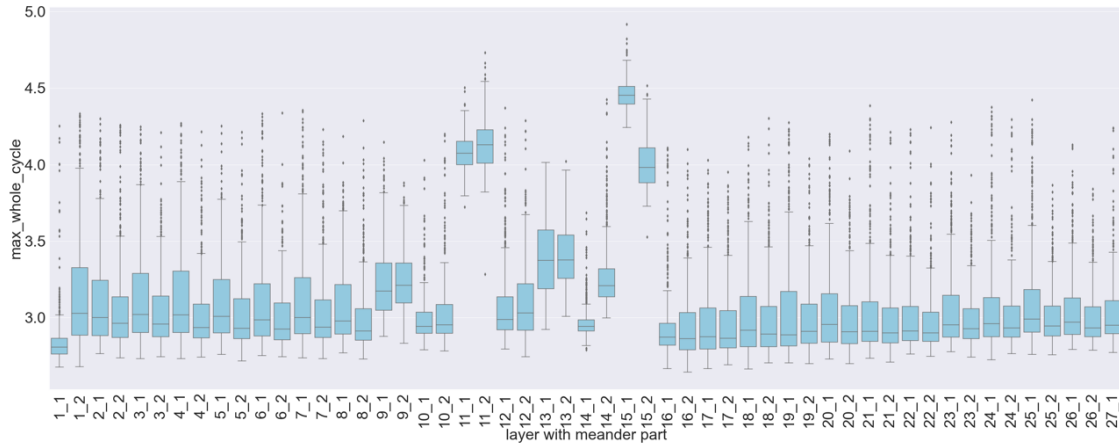


Figure 4.7: Boxplot of feature: average of maximum value of 6 consecutive voltage CMT cycles for each layer.

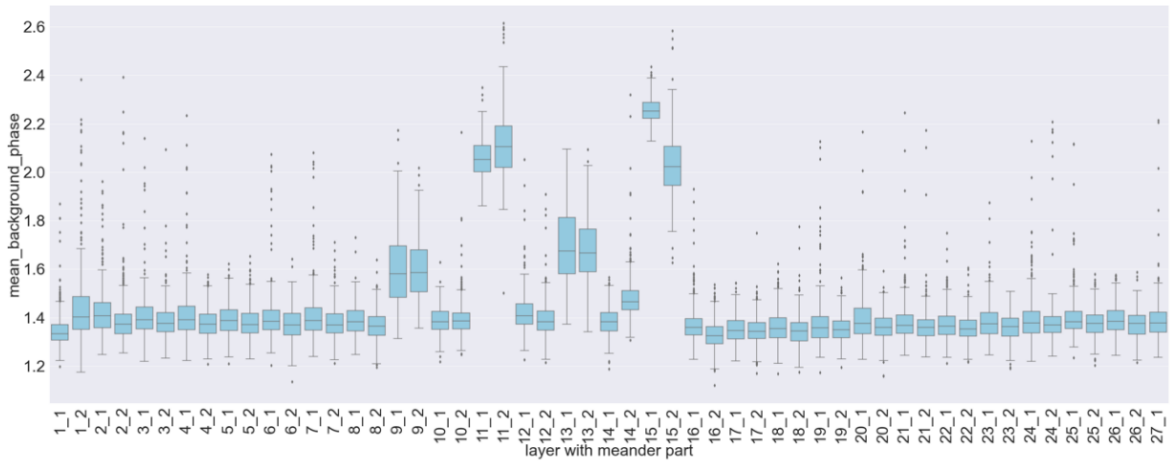


Figure 4.8: Boxplot of feature: average of mean background phase of 6 consecutive voltage CMT cycles for each layer.

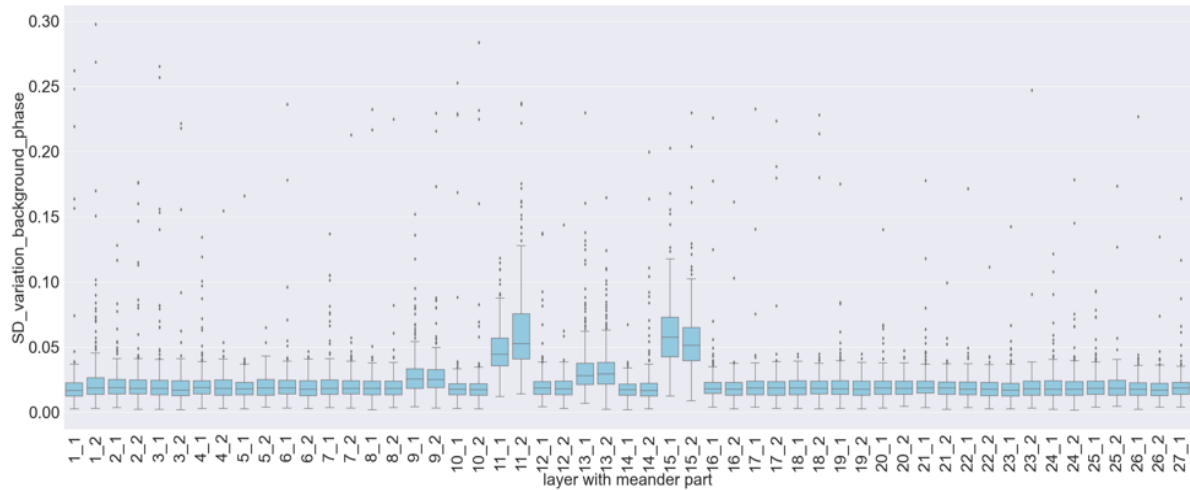


Figure 4.9: Boxplot of feature: variation of background phase of 6 consecutive voltage CMT cycles for each layer.

Model architecture is shown in the figure 4.10, dropout layers were also included in the network to reduce overfitting while training network.

Model: "sequential_1"		
Layer (type)	Output Shape	Param #
dense_4 (Dense)	(None, 24)	792
dropout_2 (Dropout)	(None, 24)	0
dense_5 (Dense)	(None, 20)	500
dense_6 (Dense)	(None, 20)	420
dropout_3 (Dropout)	(None, 20)	0
dense_7 (Dense)	(None, 32)	672
Total params: 2,384		
Trainable params: 2,384		
Non-trainable params: 0		

Figure 4.10: Architecture of developed autoencoder.

5 Methodology for Factory Data

In this chapter analysis of factory data in particular wall no. 4 is presented. As discussed in section 3.2, voltage and current CMT cycles of lab data and factory data are different. Sample rate of lab's sensor is 4000 data points per second, but factory's sensor has a sample rate of 900 data points per second. Time duration for each CMT cycle in the factory data is also less compared to data from walls no. 2 and 3 in the lab. Figures 5.1 and 5.2 shows both current and voltage cycles of factory and lab data.

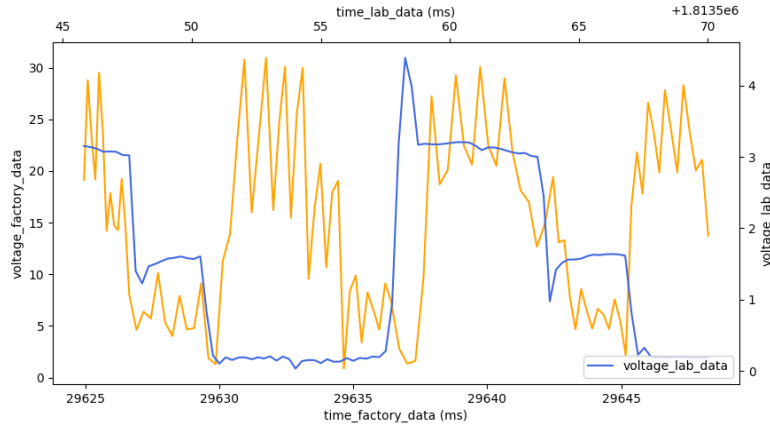


Figure 5.1: Voltage cycle of factory and lab.

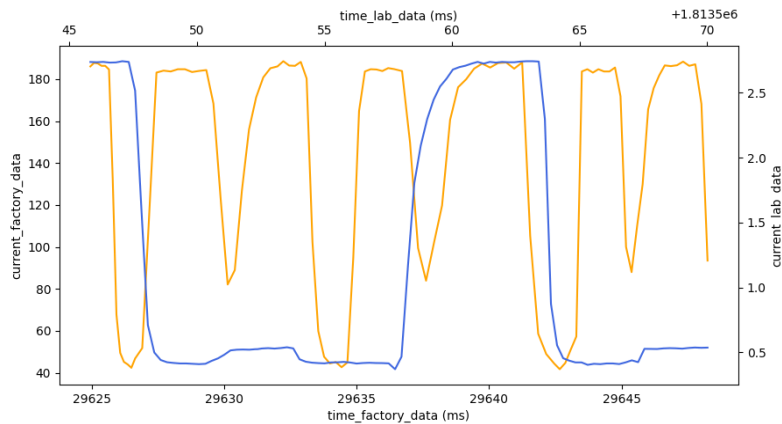


Figure 5.2: Current cycle of lab and factory data.

Voltage cycle in factory data has more fluctuation and less sample rate. To smooth the voltage cycle rolling means were applied as shown in the figure 5.3. If the window for rolling means is increased the voltage data, no longer will have parts of the cycle in different phases of CMT (refer section 2.1.1) and therefore cannot be segmented based on threshold of phase transition as discussed in section 4.2 for lab data, and if the window of the rolling mean is decreased the fluctuation in voltage value is so high that no clear phase transition is possible.

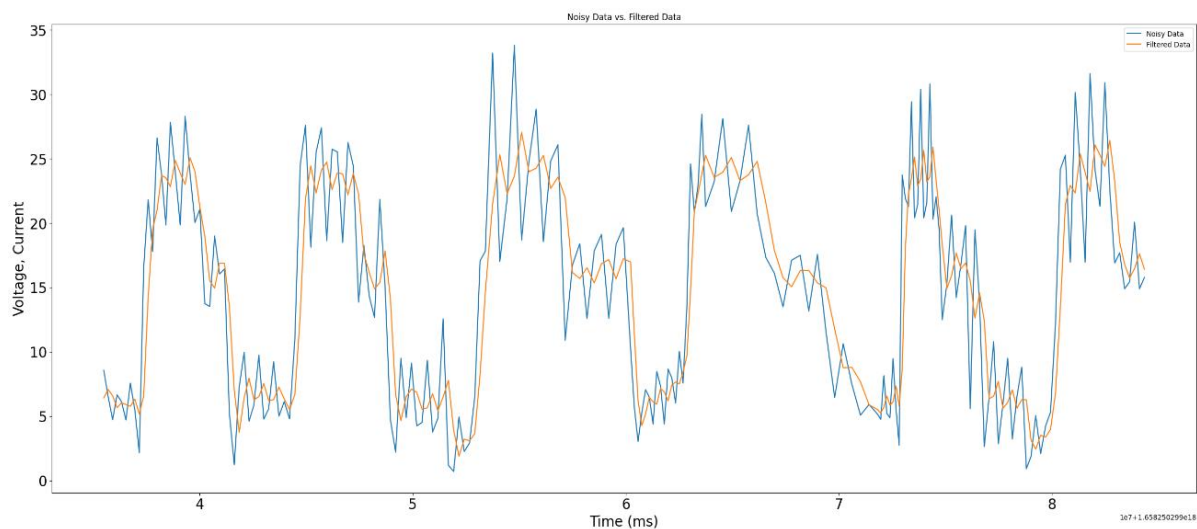


Figure 5.3: Filtered voltage data (orange).

6 Results and Discussion

This section discusses results for each of the trained models (discussed in chapter 4). All models trained for anomaly detection were with data from wall number 2, in particular, layer numbers 19th and 20th (refer section 3.5), but since current and voltage cycle are the same for wall no. 2 and wall no. 3, same models were also tested with wall no. 3 to detect anomalies and check transferability. Results for each model for both walls 2 and 3 are presented in this chapter. This chapter also discusses how developed models for lab data can be transferred to factory data.

6.1 Result For Model Based on Forecasting In Unsupervised Approach

Forecasting-based model, Inception Time (refer section 2.7.3) was not able to classify anomalous data in reduced gas flow regions in wall no. 2 as anomalous. Threshold of forecasting error i.e., mean squared error between the forecasted and original signal to classify as anomalous was changed by changing the amount of contamination in training data but model failed to correctly classify the data from reduced gas flow regions as anomalous. This approach was not further used to test on wall no. 3 because of unreliable results for oxidized regions in wall no. 2.

6.2 Result For Model Based on Single CMT Cycle Using Unsupervised Learning

In this section, results for both reconstruction-based and feature-based models discussed in section 4.2 are presented.

6.2.1 Reconstruction-based learning

Quantitative and qualitative results of model's performance for both wall no. 2 and wall no. 3 are presented below. Amount of contamination in normal training data was set manually by looking at the true positive rate and defect detection ability of the model.

Second row in each table represents the results when 10mm on the sides of walls were not analyzed because sides of the layers had too many model anomalies with no correlation to defects.

Wall no. 2 results: Table 6.1 shows amount of contamination set in training data, anomaly threshold with respect to contamination amount, true positive rate, and percentage of defect detected based on volume by the model when tested on wall no. 2. Results do not include the region where gas flow was reduced.

Row no.	Contamination % for training data	Threshold for anomaly score	True positive rate	Detection % low volume defect (volume less than 1 mm ³)	Detection % medium volume defect (volume between 1 mm ³ to 5 mm ³)	Detection % for high volume defects (volume from 5 mm ³ onwards)
1	3.2 %	0.000192	26 %	100 %	100 %	100 %
2	3.2 %	0.000192	44.37 %	94.44 %	100 %	100 %

Table 6.1: Quantitative results for wall no. 2 with reconstruction-based model using model evaluation algorithm.

Figure 6.1 shows anomalies and normal data detected by the model overlaid on a CT scan picture for 5th layer with unintentional small pores. Picture below (figure 6.2) shows correlation of model anomalies and defects for 5th layer by using evaluation algorithm (discussed in section 4.1). Figure 6.3 shows that model detected anomalies robustly in data of layer 11 where gas flow was reduced.

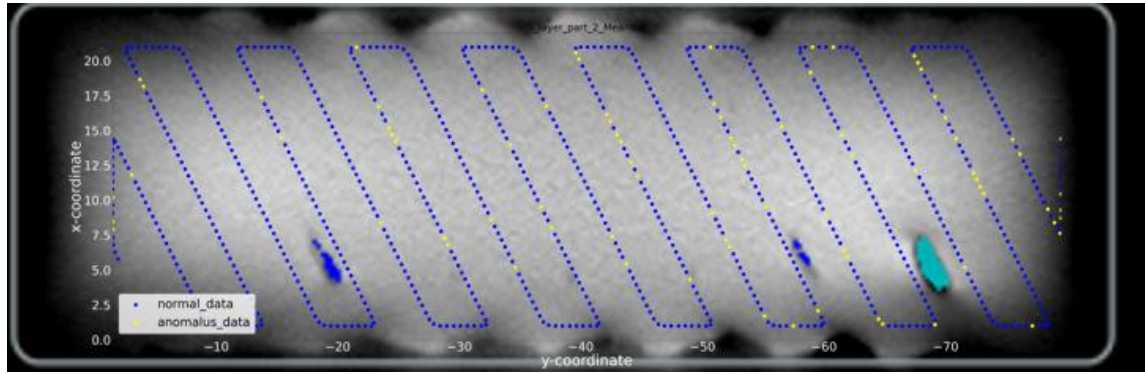


Figure 6.1: Shows layer no. 5th's XY cross-sectional view of CT scan with low volume defects (colored) overlaid with model's results.

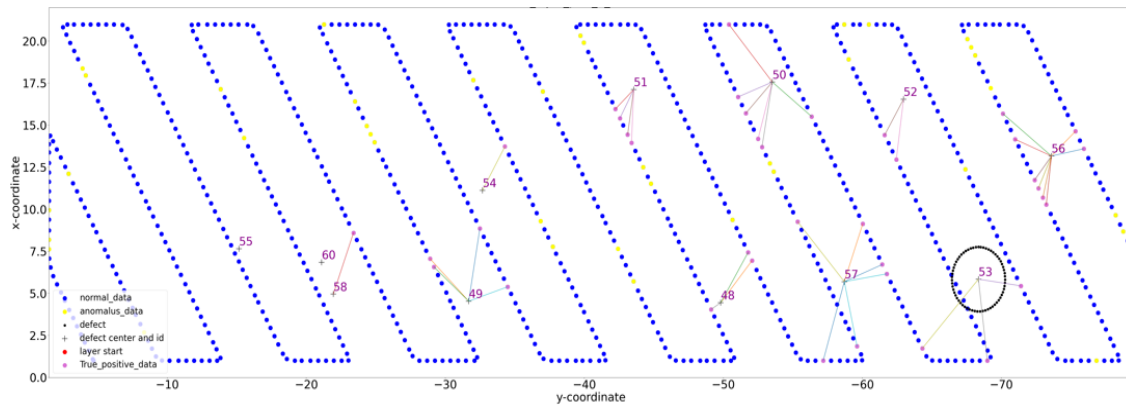


Figure 6.2: Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 5th.

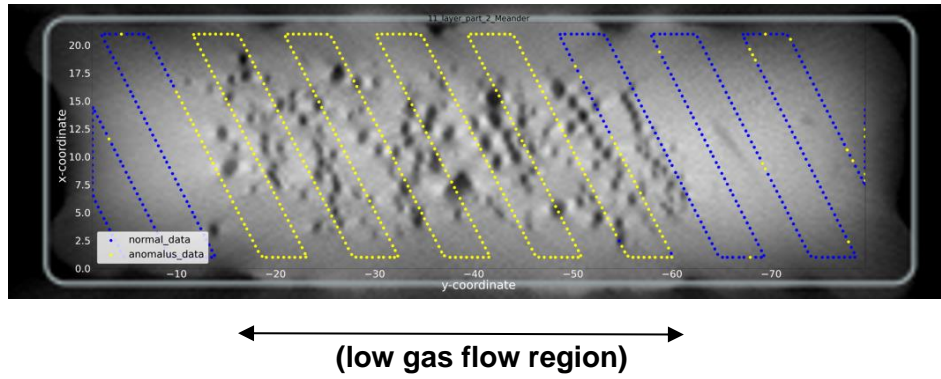


Figure 6.3: Anomalous and normal data detected by model in layer 11 with low gas flow (oxidized) overlaid over a CT scan picture of the layer.

Wall 3 results: Table 6.2 shows amount of contamination set in training data, anomaly threshold with respect to contamination amount, true positive rate, and percentage of defect detected based on volume by the model when tested on wall 3.

Row no.	Contamination % for training data	Threshold for anomaly score	True positive rate	Detection % low volume defect (volume less than 1 mm ³)	Detection % medium volume defect (volume between 1 mm ³ to 50 mm ³)	Detection % for high volume defects (volume from 50 mm ³ onwards)
1	3.2 %	0.000192	30 %	84.52 %	100 %	100 %
2	3.2 %	0.000192	48.27 %	86.80 %	100 %	100 %

Table 6.2: Quantitative results for wall no. 3 with reconstruction-based model using model evaluation algorithm.

13th layer (figure 6.4) has large number of continuous anomalous cycles (circled in orange) near large cavity suggesting that this defect correlates with model anomalies. Figure 6.5 shows a correlation between model anomalies and defects for 13th layer by using the evaluation algorithm, since the information given for the global center of defect is used to project the defect on the layer the evaluation algorithm fails to consider all continuous model anomalies (in orange circle) as true positive.

Layer number 12th with large and small defects along with the correlation between model anomalies and defects, using the evaluation algorithm is shown in figures 6.6 and 6.7.

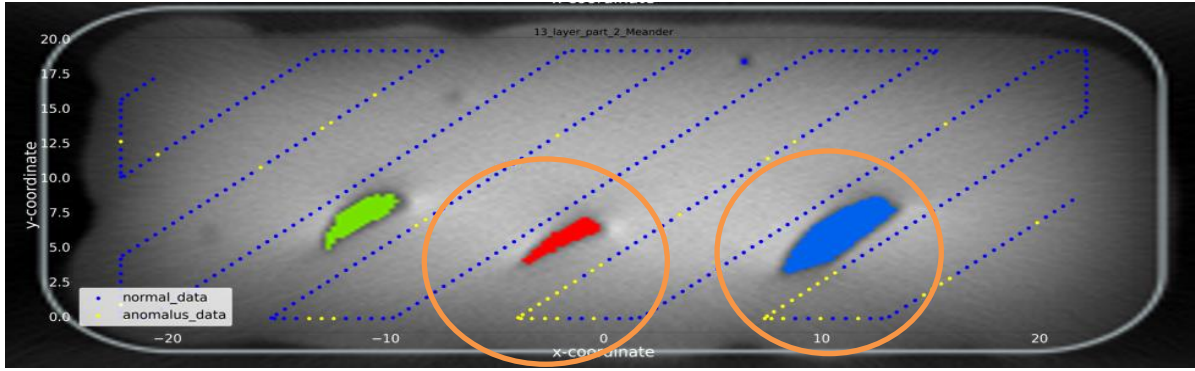


Figure 6.4: Shows layer no. 13th XY cross-sectional view of CT scan with large defects (colored) overlayed with model's results.

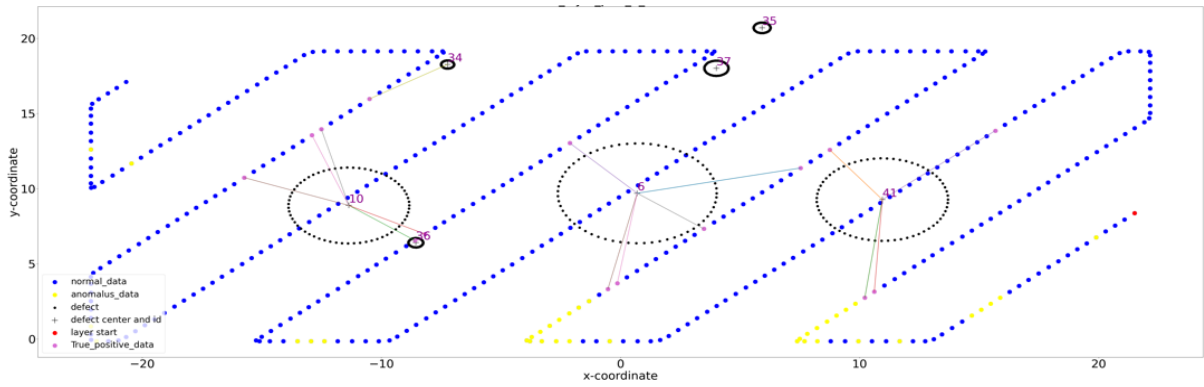


Figure 6.5: Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 13th.

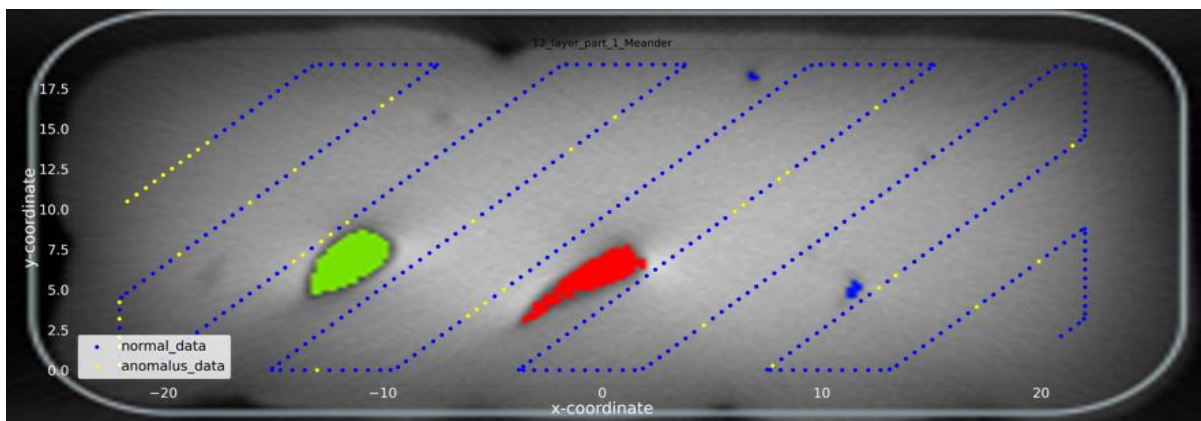


Figure 6.6: Shows layer no. 12th XY cross-sectional view of CT scan with large defects (colored) overlayed with model's results.

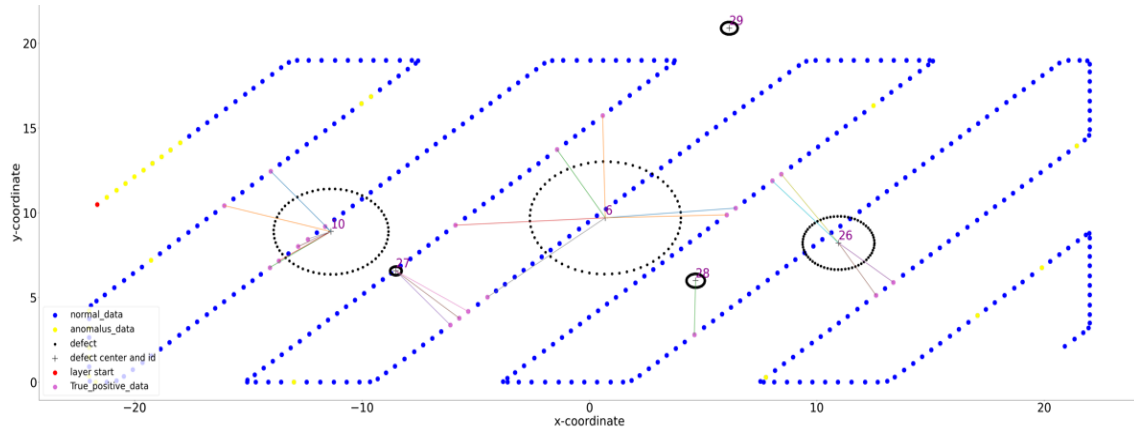


Figure 6.7: Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 12.

Discussion: Model was able to identify oxidized regions very accurately, it also has anomaly detection capability for both pores and large cavities but the true positive rate is low, meaning some of the anomalies in voltage data were random and not because of a real defect. Because model analyzes a single CMT cycle, there is a possibility that the CMT cycle is anomalous but does not result in a defect, and thus is counted as a false positive.

6.2.2 Feature-based learning

Quantitative and qualitative results of model's performance for both wall no. 2 and wall no. 3 are presented below. Amount of contamination in normal training data was set manually by looking at the true positive rate and defect detection ability of the model.

Wall 2 results: Table 6.3 shows amount of contamination set in training data, anomaly threshold with respect to contamination amount, true positive rate, and percentage of defect detected based on volume by the model when tested on wall 2. Results do not include the region where gas flow was reduced.

Row no.	Contamination % for training data	Threshold for anomaly score	True positive rate	Detection % low volume defect (volume less than 1 mm ³)	Detection % medium volume defect (volume between 1 mm ³ to 5 mm ³)	Detection % for high volume defects (volume from 5 mm ³ onwards)
1	3.2 %	1.0199	21.14 %	96.52 %	100 %	100 %
2	3.2 %	1.0199	44.15 %	93.51 %	100 %	100 %

Table 6.3: Quantitative results for wall no. 2 with feature-based model using model evaluation algorithm.

Figure 6.8 shows model results overlayed on a CT scan picture for 5th layer with unintentional small pores. Figure 6.9 shows that model detected anomalies robustly in the data of layer 11 where gas flow was reduced to induce oxidation.

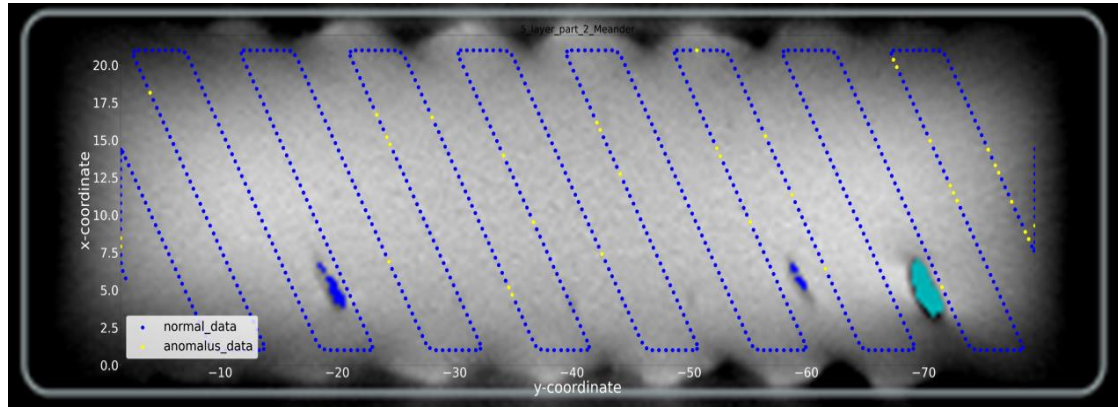


Figure 6.8: Shows layer no. 5th's XY cross sectional view of CT scan with low volume defects (colored) overlayed with model's results.

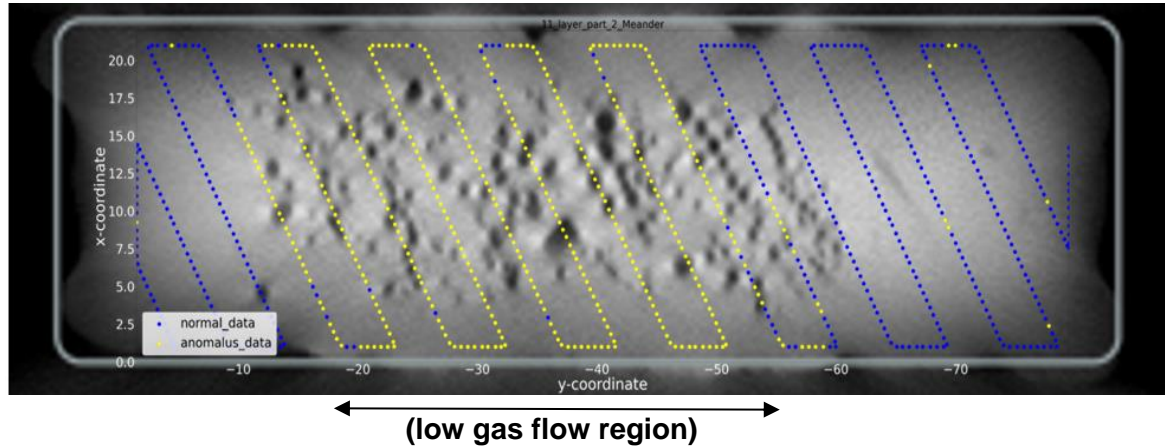


Figure 6.9: Anomalies and normal data detected by model in layer 11 with low gas flow (oxidized) overlayed over CT scan of the layer.

Wall 3 results: Table 6.4 shows amount of contamination set in training data, anomaly threshold with respect to contamination amount, true positive rate, and percentage of defect detected based on volume by the model when tested on wall 3.

Row no.	Contamination % for training data	Threshold for anomaly score	True positive rate	Detection % low volume defect (volume less than 1 mm ³)	Detection % medium volume defect (volume between 1 mm ³ to 50 mm ³)	Detection % for high volume defects (volume from 50 mm ³ onwards)
1	3.2 %	1.0199	33.65 %	98.80 %	100 %	100 %
2	3.2 %	1.0199	50.65 %	98.68 %	100 %	100 %

Table 6.4: Quantitative results for wall no. 3 with feature-based model using model evaluation algorithm.

Although model performed well when compared to reconstruction based quantitatively, the figures 6.10 and 6.11 shows that results of this model gave lots of anomalous cycles spread throughout the layer. This suggests that feature-based method for a single cycle does not yield reliable results.

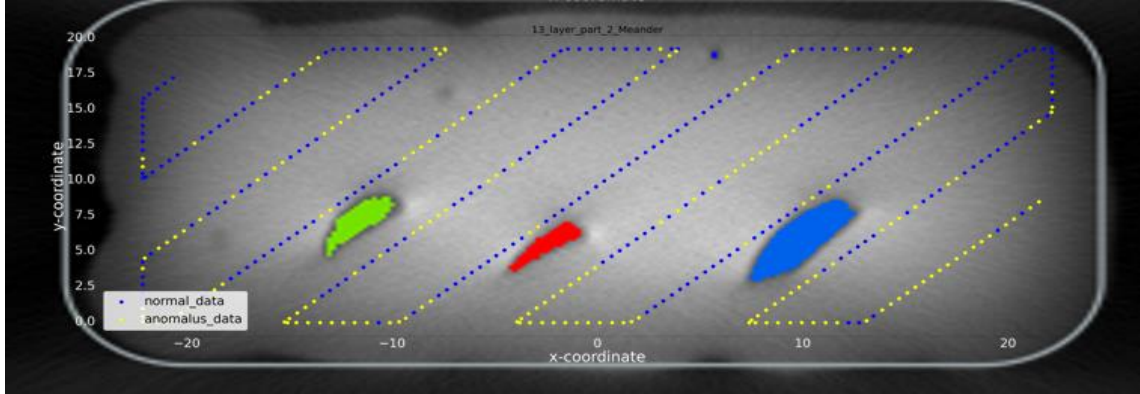


Figure 6.10: Shows layer no. 13th XY cross-sectional view of CT scan with large defects (colored) overlaid with model's results.

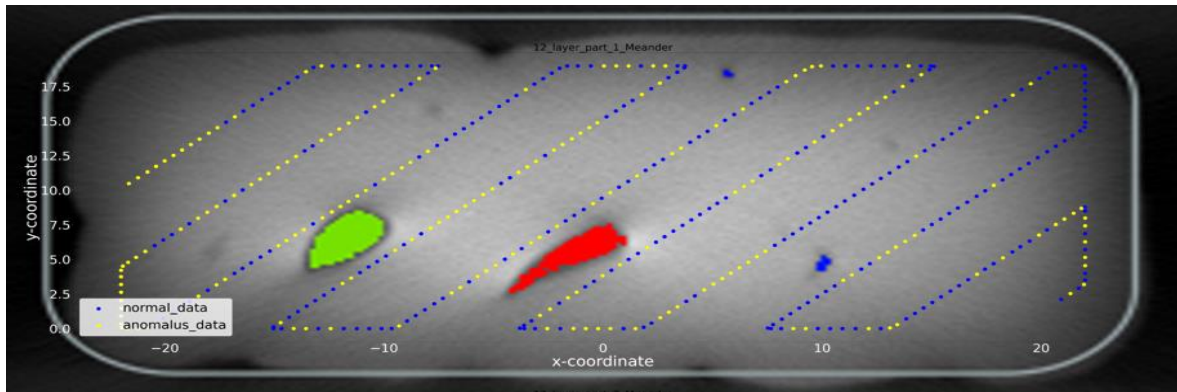


Figure 6.11: Shows layer no. 12th XY cross-sectional view of CT scan with large and small defects (colored) overlaid with model's results.

Discussion: Model was able to identify oxidized regions very accurately. Model outputs many anomalous cycles as can be seen in figures 6.10 and 6.11 and is therefore unreliable. Moreover, the assumption that a single CMT cycle is sufficient to induce a true defect may not be correct, and considering the number of continuous anomalous data points as a group leading to a defect is likely correct, but number of continuous anomalous data points to consider must be set manually. No further analysis for a single CMT cycle was done because setting how many anomalous cycles together in the vicinity can produce defects was unclear.

6.3 Result For Model Based on Multiple CMT Cycle Using Unsupervised Learning

Quantitative and qualitative results of model's performance for both wall no. 2 and wall no. 3 are presented below. Amount of contamination in normal training data was set manually by looking at the true positive rate and defect detection ability of the model.

Wall 2 results: Table 6.5 shows amount of contamination set, anomaly threshold with respect to contamination amount, true positive rate, and percentage of defect detected based on volume by the model when tested on wall 2. Results do not include the region where gas flow was reduced.

Row no.	Contamination % for training data	Threshold for anomaly score	True positive rate	Detection % low volume defect (volume less than 1 mm ³)	Detection % medium volume defect (volume between 1 mm ³ to 5 mm ³)	Detection % for high volume defects (volume from 5 mm ³ onwards)
1	6.4 %	0.735	24.55 %	79.13 %	100 %	87.39 %
2	6.4 %	0.735	43.58	78.70	100 %	90.17 %

Table 6.5: Quantitative results for wall no. 2 with trained model using model evaluation algorithm.

Figure 6.12 shows model results overlayed on a CT scan picture for 5th layer with unintentional small pores. Figure 6.13 shows correlation of model anomalies and defects for 5th layer by using evaluation algorithm. Although quantitatively the model performs comparable to other models but qualitatively the model has superior performance because of more temporal and spatial information. Figure 6.14 shows that model detected anomalies robustly in data of layer 11 where gas flow was reduced. Principal component analysis of features from voltage data for low gas flow region and normal training data is shown in figure 6.15. Principal components of features of low gas flow regions and normal training data have isolated clusters and therefore oxidation can be reliably detected.

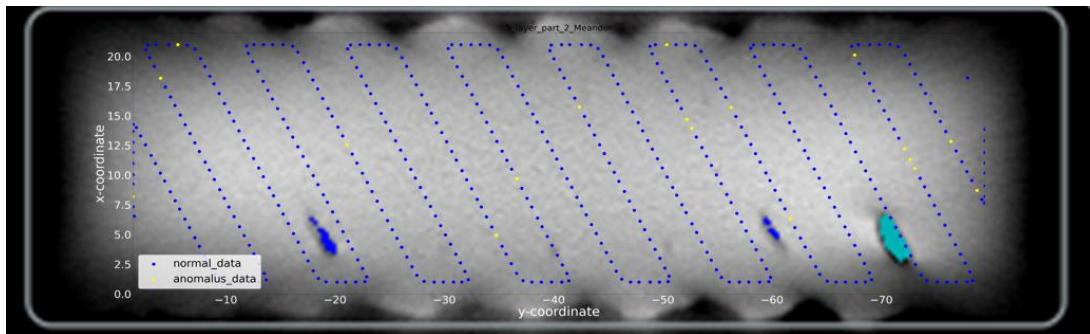


Figure 6.12: Shows layer no. 5th's XY cross sectional view of CT scan with low volume defects (colored) overlayed with model's results.

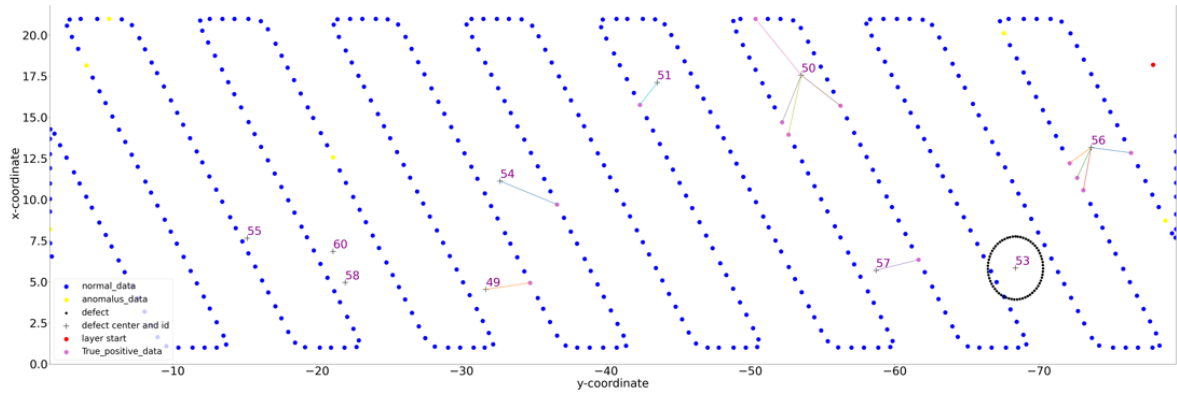


Figure 6.13: Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 5th.

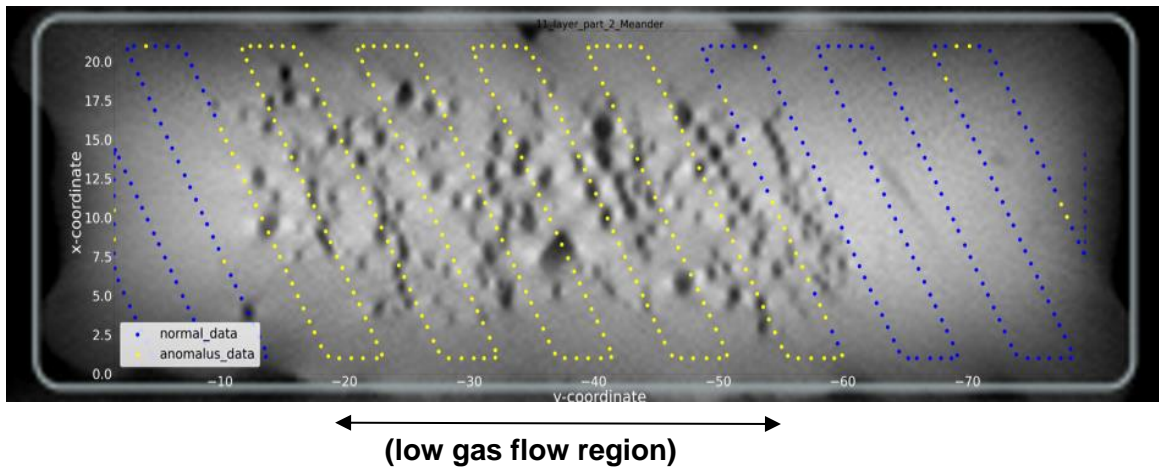


Figure 6.14: Shows anomalies and normal data detected by model in layer 11 with low gas flow (oxidized) overlaid over CT scan of the layer.

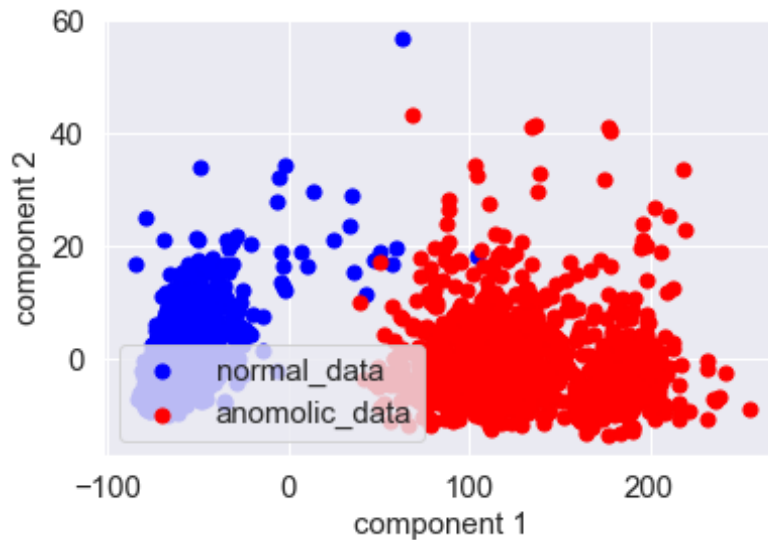


Figure 6.15: Principal components of voltage data (comprising features of 6 consecutive cycles) of region with low gas flow (defective) and normal training data (as discussed in section 3.5).

Wall 3 results: Table 6.6 shows amount of contamination set in training data, anomaly threshold with respect to contamination amount, true positive rate, and percentage of defect detected based on volume by the model when tested on wall 3.

Row no.	Contamination % for training data	Threshold for anomaly score	True positive rate	Detection % low volume defect (volume less than 1 mm ³)	Detection % medium volume defect (volume between 1 mm ³ to 50 mm ³)	Detection % for high volume defects (volume from 50 mm ³ onwards)
1	6.4 %	0.735	35.89 %	73.80 %	100 %	100 %
2	6.4 %	0.735	52.48 %	71.05 %	100 %	100 %

Table 6.6: Quantitative results for wall no. 3 with trained model using model evaluation algorithm.

Layer 13th (figure 6.16) has large number of continuous anomalous cycles near a large cavity suggesting that these large cavities correlate with model anomalies. Figure 6.17 shows correlation between model anomalies and defects for layer 13th using evaluation algorithm. Layer number 12th with large and small gaps is shown below in figure 6.18 along with the correlation between model anomalies and defects, using evaluation algorithm in figure 6.19. As can be seen from the comparison of figures 6.16 and 6.17 that due to the projection of global center of defect many true positive anomalous data (in orange circle) were considered false positive resulting in overall bad performance when calculated using evaluation algorithm.

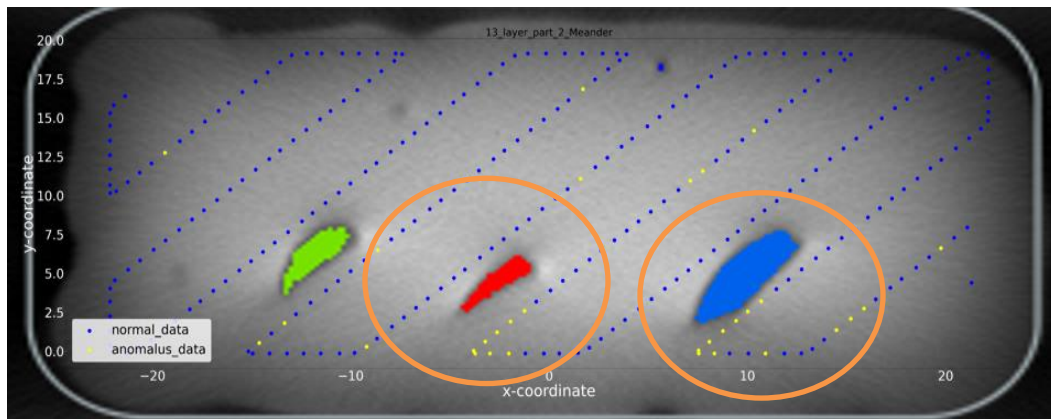


Figure 6.16: Shows layer no. 13th XY cross-sectional view of CT scan with large defects (colored) overlayed with model's result.

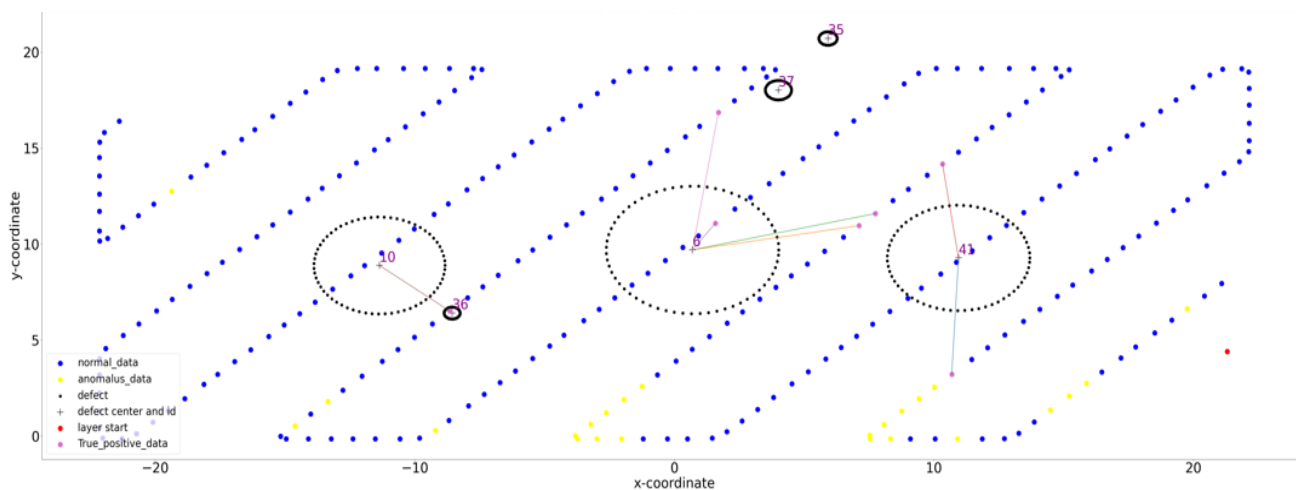


Figure 6.17: Correlation found between model anomalies and defects (joined by lines) by using evaluation algorithm for layer 13th.

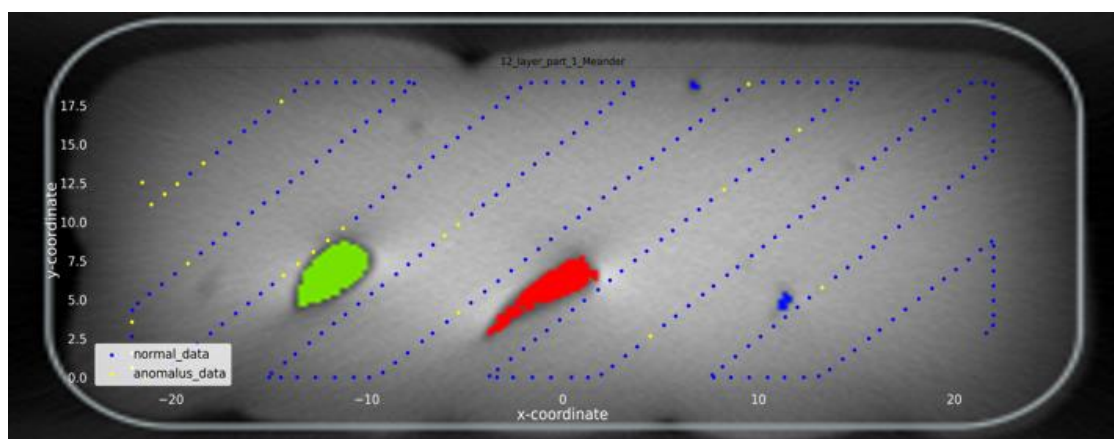


Figure 6.18: Shows layer no. 12th XY cross-sectional view of CT scan with large defects (colored) overlaid with model's results.

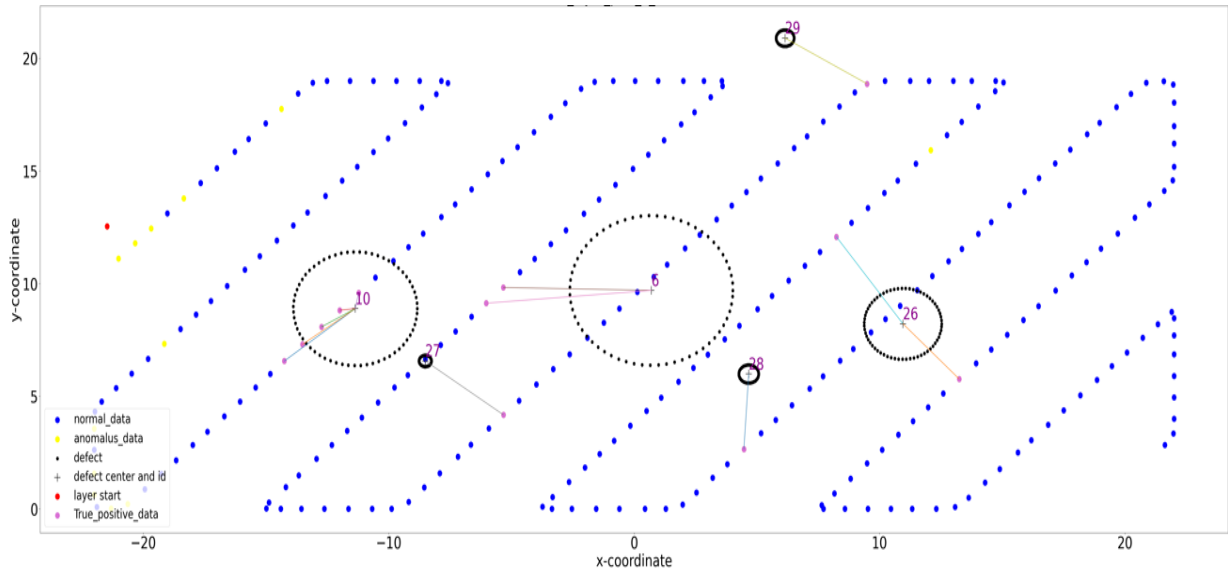


Figure 6.19: Correlation found between model anomalies and defect (joined by lines) by using evaluation algorithm for layer 12.

Discussion: Model can reliably predict anomalies in oxidized regions. Model has good defect detection capability but the true positive rate is still low, meaning not all model anomalies lead to defects and therefore the anomaly detection model is not accurate, partly due to uncertainty in consideration of the center for large defects the quantitative results discussed in table 6.6 are bad. Evaluation of 6 consecutive cycles has advantage of more temporal and spatial information. Model has a higher true positive rate when compared to the reconstruction-based method, and is the most reliable.

6.4 Transferability On Factory Data

Models developed for lab data could not be applied directly to factory data because of different voltage and current cycles. A new algorithm to detect the phases and features of factory data should be developed and then the transferability of the models developed for lab data can be checked with factory data (production).

7 Conclusions and Future Work

This chapter concludes the thesis by summarizing the issues addressed and relevant insights gained throughout the process. Finally, several potential future work ideas are addressed.

7.1 Conclusion

Goal of the thesis was development of anomaly detection system for Wire arc additive manufacturing (WAAM) process with steel alloy as material using voltage data. Several approaches for model development were tried and evaluation algorithm for model performance calculation was developed.

The major problem of unlabeled data was addressed by using unsupervised machine learning approaches. Models were trained in an unsupervised fashion by taking data from almost defect free layers from oxidized complex walls i.e., wall no 2, and same models were tested with wall no. 3 (with similar process parameters as wall no. 2), the models were able to detect anomalies in both walls. The thesis approached the problem with different methods and gives insights of which method is suitable to detect anomalies reliably. Convolutional network and autoencoder were optimized and trained to learn non-anomalous data patterns. Moreover, thesis also highlighted the relationship between different features of voltage data with defects. Quantitative and qualitative results of each model were also presented in the work to give more insight into performance of the models. Models were able to detect anomalies in voltage data during oxidation, near pores and, large cavities indicating that detected anomalies led to the defect formation, although models are not reliable for detecting formation of pores and cavities. Model with input of features of 6 CMT cycles performed the best. Thesis also compared voltage and current data in Siemens lab and factory (production) and discussed why the developed models cannot directly be transferred for use with factory data.

7.2 Future work

Different approaches can be used to further make the model robust. Due to uncertainty in defect locations some of the true positive anomalous data correlating to defect are considered false positive, to deal with this issue, a rule could be implemented in the model evaluation algorithm (section 4.1) that if any anomalous data correlates with defect, then all continuous anomalous data near it will be true positive. All the true positive data can be labeled as anomalous and a separate neural network can be trained in a supervised manner, and the same network can be used as anomaly detection model. Generative models could also be used to further create labeled anomalous data. The assumption of defects being circular in the model evaluation algorithm is inaccurate, defect generally has the shape of ellipse therefore the evaluation algorithm can be updated to consider defect as an ellipse rather than circle shaped.

As discussed in thesis, that lab and factory data are different and therefore raw voltage signal-based approaches like forecasting can be further investigated with different inputs and output

numbers. Using transfer learning the already trained model for lab data could be used for new kinds of data. Number of features for feature-based approaches could be reduced further by looking at feature importance and inter-correlation between the features, this will reduce model latency. Finally, anomaly detection from other sensors like audio or camera can be used to get confidence on the prediction by voltage data-based anomaly detection system.

8 Bibliography

- [Ahr + 19] Ahrens, L., Ahrens, J., & Schotten, H. D. (2019). A machine-learning phase classification scheme for anomaly detection in signals with periodic characteristics. *EURASIP Journal on Advances in Signal Processing*, 2019(1). <https://doi.org/10.1186/s13634-019-0619-3>
- [Ban + 20] Bank, D., Koenigstein, N., & Giryas, R. (2020). Autoencoders. *ArXiv: Learning*. <https://doi.org/10.48550/arXiv.2003.05991>
- [Chr + 18] Christ, M., Braun, N., Neuffer, J., & Kempa-Liehr, A. W. (2018). Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). *Neurocomputing*, 307, 72–77. <https://doi.org/10.1016/j.neucom.2018.03.067>
- [Cun + 17] Cunningham, C., Wikshåland, S., Xu, F., Kemakolam, N., Shokrani, A., Dhokia, V., & Newman, S. (2017). Cost Modelling and Sensitivity Analysis of Wire and Arc Additive Manufacturing. *Procedia Manufacturing*, 11, 650–657. <https://doi.org/10.1016/j.promfg.2017.07.163>
- [Eli + 18] Elie Aljalbout, Vladimir Golkov, Yawar Siddiqui, & Daniel Cremers. (2018). Clustering with Deep Learning: Taxonomy and New Methods. *ArXiv: Learning*. <https://arxiv.org/pdf/1801.07648.pdf>
- [Fel +19] Feldmann, M., Kühne, R., Citarelli, S., Reisgen, U., Sharma, R., & Oster, L. (2019). 3D-Drucken im Stahlbau mit dem automatisierten Wire Arc Additive Manufacturing. *Stahlbau*, 88(3), 203–213. <https://doi.org/10.1002/stab.201800029>
- [Ful + 18] Fulcher, B. D. (2018). Feature-Based Time-Series Analysis. *Feature Engineering for Machine Learning and Data Analytics*, 87–116. <https://doi.org/10.1201/9781315181080-4>
- [Hau + 21a] Hauser, T., Reisch, R. T., Breese, P. P., Nalam, Y., Joshi, K. S., Bela, K., Kamps, T., Volpp, J., & Kaplan, A. F. (2021). Oxidation in wire arc additive manufacturing of aluminium alloys. *Additive Manufacturing*, 41, 101958. <https://doi.org/10.1016/j.addma.2021.101958>
- [Hau + 21b] Hauser, T., Reisch, R. T., Breese, P. P., Lutz, B. S., Pantano, M., Nalam, Y., Bela, K., Kamps, T., Volpp, J., & Kaplan, A. F. (2021). Porosity in wire arc additive manufacturing of aluminium alloys. *Additive Manufacturing*, 41, 101993. <https://doi.org/10.1016/j.addma.2021.101993>

- [Hua + 20] Huang, Y., Yuan, Y., Yang, L., Wu, D., & Chen, S. (2020). Real-time monitoring and control of porosity defects during arc welding of aluminum alloys. *Journal of Materials Processing Technology*, 286, 116832. <https://doi.org/10.1016/j.jmatprotec.2020.116832>
- [Ism + 20] Ismail Fawaz, H., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F., Weber, J., Webb, G. I., Idoumghar, L., Muller, P. A., & Petitjean, F. (2020). InceptionTime: Finding AlexNet for time series classification. *Data Mining and Knowledge Discovery*, 34(6), 1936–1962. <https://doi.org/10.1007/s10618-020-00710-y>
- [Jin + 20] Jin, W., Zhang, C., Jin, S., Tian, Y., Wellmann, D., & Liu, W. (2020). Wire Arc Additive Manufacturing of Stainless Steels: A Review. *Applied Sciences*, 10(5), 1563. <https://doi.org/10.3390/app10051563>
- [Li + 19] Li, J. Z., Alkahari, M. R., Rosli, N. A. B., Hasan, R., Sudin, M. N., & Ramli, F. R. (2019). Review of Wire Arc Additive Manufacturing for 3D Metal Printing. *International Journal of Automation Technology*, 13(3), 346–353. <https://doi.org/10.20965/ijat.2019.p0346>
- [Moh + 20] Mohammad Braei, & Sebastian Wagner. (2020). Anomaly Detection in Univariate Time-series: A Survey on the State-of-the-Art. *ArXiv: Learning*. <http://export.arxiv.org/pdf/2004.00433>
- [Mun + 19] Munir, M., Siddiqui, S. A., Dengel, A., & Ahmed, S. (2019). DeepAnT: A Deep Learning Approach for Unsupervised Anomaly Detection in Time Series. *IEEE Access*, 7, 1991–2005. <https://doi.org/10.1109/access.2018.2886457>
- [Ngu + 19] Nguyen, V. (2019). Bayesian Optimization for Accelerating Hyper-Parameter Tuning. 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE). <https://doi.org/10.1109/aike.2019.00060>
- [M16] Nielsen, M. (2016). *Neural Networks and Deep Learning*. Determination press. <http://neuralnetworksanddeeplearning.com/>
- [Mal + 16] Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). LSTM-based Encoder-Decoder for Multi-sensor Anomaly Detection. *ArXiv: Artificial Intelligence*. <https://arxiv.org/pdf/1607.00148.pdf>

- [myv] MyVgl. A myVGL. (2020). Industrial CT Software | Volume Graphics. <https://www.volumegraphics.com/en/products/myvgl.html>
- [Qi + 19] Qi, X., Chen, G., Li, Y., Cheng, X., & Li, C. (2019). Applying Neural-Network-Based Machine Learning to Additive Manufacturing: Current Applications, Challenges, and Future Perspectives. *Engineering*, 5(4), 721–729. <https://doi.org/10.1016/j.eng.2019.04.012>
- [Pla + 19] Plangger J, Schabhüttl P, Vuherer T, Enzinger N. CMT Additive Manufacturing of a High Strength Steel Alloy for Application in Crane Construction. *Metals*. 2019; 9(6):650. <https://doi.org/10.3390/met9060650>
- [Cha + 19] Chalapathy R., & Sanjay Chawla. (2019). Deep Learning for Anomaly Detection: A Survey. *ArXiv: Learning*. <https://arxiv.org/pdf/1901.03407v2.pdf>
- [Rei + 20a] Reisch, R., Hauser, T., Lutz, B., Pantano, M., Kamps, T., & Knoll, A. (2020). Distance-Based Multivariate Anomaly Detection in Wire Arc Additive Manufacturing. 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA). <https://doi.org/10.1109/icmla51294.2020.00109>
- [Rei + 20b] Reisch, R., Hauser, T., Kamps, T., & Knoll, A. (2020). Robot Based Wire Arc Additive Manufacturing System with Context-Sensitive Multivariate Monitoring Framework. *Procedia Manufacturing*, 51, 732–739. <https://doi.org/10.1016/j.promfg.2020.10.103>
- [Cho + 20] Rene Y. Choi, Aaron S. Coyner, Jayashree Kalpathy-Cramer, Michael F. Chiang, & J. Peter Campbell. (2020). Introduction to Machine Learning, Neural Networks, and Deep Learning. *Translational Vision Science & Technology*, 9(2), 14. <https://doi.org/10.1167/tvst.9.2.14>
- [Rod +19] Rodrigues, T. A., Duarte, V., Miranda, R., Santos, T. G., & Oliveira, J. (2019). Current Status and Perspectives on Wire and Arc Additive Manufacturing (WAAM). *Materials*, 12(7), 1121. <https://doi.org/10.3390/ma12071121>
- [Rud + 16] Ruder, S. (2016). An overview of gradient descent optimization algorithms. *ArXiv: Learning*. <https://doi.org/10.48550/arXiv.1609.04747>
- [Sel + 18] Selvi, S., Vishvaksenan, A., & Rajasekar, E. (2018). Cold metal transfer (CMT) technology - An overview. *Defence Technology*, 14(1), 28–44. <https://doi.org/10.1016/j.dt.2017.08.002>

- [SV14] Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding Machine Learning: From Theory to Algorithms (1st ed.). Cambridge University Press.
- [Shi + 20] Shin, S., Jin, C., Yu, J., & Rhee, S. (2020). Real-Time Detection of Weld Defects for Automated Welding Process Base on Deep Neural Network. *Metals*, 10(3), 389. <https://doi.org/10.3390/met10030389>
- [BLL10] Siciliano, B., Sciavicco, L., Villani, L., & Oriolo, G. (2010). *Robotics: Modelling, Planning and Control* (Advanced Textbooks in Control and Signal Processing) (Softcover reprint of hardcover 1st ed. 2009). Springer. <https://link.springer.com/book/10.1007/978-1-84628-642-1>
- [Sim + 18] Simeone, O. (2018). A Very Brief Introduction to Machine Learning With Applications to Communication Systems. *IEEE Transactions on Cognitive Communications and Networking*, 4(4), 648–664. <https://doi.org/10.1109/tccn.2018.2881442>
- [Tan + 19] Tan, Y., Jin, B., Nettekoven, A., Chen, Y., Yue, Y., Topcu, U., & Sangiovanni-Vincentelli, A. (2019). An Encoder-Decoder Based Approach for Anomaly Detection with Application in Additive Manufacturing. 2019 18th IEEE International Conference on Machine Learning and Applications (ICMLA). <https://doi.org/10.1109/icmla.2019.00171>
- [Tsa + 2022] tsai. (2022). A state-of-the-art deep learning library for time series and sequential data <https://timeseriesai.github.io/tsai/>
- [Wes + 10] Wesling, V., Schram, A., & Kessler, M. (2010). Low Heat Joining – Manufacturing and Fatigue Strength of Brazed, Locally Hardened Structures. *Advanced Materials Research*, 137, 347–374. <https://doi.org/10.4028/www.scientific.net/amr.137.347>
- [Wil + 16] Williams, S. W., Martina, F., Addison, A. C., Ding, J., Pardal, G., & Colegrove, P. (2016). Wire + Arc Additive Manufacturing. *Materials Science and Technology*, 32(7), 641–647. <https://doi.org/10.1179/1743284715y.0000000073>
- [Wu + 18] Wu, B., Pan, Z., Ding, D., Cuiuri, D., Li, H., Xu, J., & Norrish, J. (2018). A review of the wire arc additive manufacturing of metals: properties, defects and quality improvement. *Journal of Manufacturing Processes*, 35, 127–139. <https://doi.org/10.1016/j.jmapro.2018.08.001>

- [Xu + 18] Xu, F., Dhokia, V., Colegrove, P., McAndrew, A., Williams, S., Henstridge, A., & Newman, S. T. (2018). Realisation of a multi-sensor framework for process monitoring of the wire arc additive manufacturing in producing Ti-6Al-4V parts. *International Journal of Computer Integrated Manufacturing*, 31(8), 785–798. <https://doi.org/10.1080/0951192x.2018.1466395>
- [Guo + 18] Guo Y., Liao W., Wang Q., Yu L., Ji T., & Li P.. (2018). Multidimensional Time Series Anomaly Detection: A GRU-based Gaussian Mixture Variational Autoencoder Approach. *Asian Conference on Machine Learning*, 97–112. Retrieved from <https://proceedings.mlr.press/v95/guo18a/guo18a.pdf>