NAZARBAYEV UNIVERSITY

NAZARBAYEV UNIVERSITY

SCHOOL OF ENGINEERING AND DIGITAL SCIENCES

ENG 400 - Capstone Project

Machine Learning in Digital Twin for Detection and Prediction of Defects in Metal 3D Printing

Group Members

Bakbergen Yermekbayev

Dinmukhammed Azamatov

Ali Yelshibek

Askhat Manapov

Supervisors: Professor Essam Shehab & Associate Professor Md.Hazrat Ali

Declaration

We hereby declare that this report entitled "Machine Learning in Digital Twin for Detection and Prediction of Defects in Metal 3D Printing" is the result of our project work except for quotations and citations that have been duly acknowledged. We also declare that it has not been previously or concurrently submitted for any other degree at Nazarbayev University.

Names:

Bakbergen Yermekbayev	201926481
Dinmukhammed Azamatov	201910741
Ali Yelshibek	201965308
Askhat Manapov	201999187

Date: April 27, 2025

Abstract

The Directed Energy Deposition (DED) 3D printing process has emerged as an innovative technology for manufacturing complex metal components in today's industries. However, its broader industrial adoption is limited by challenges such as defects in printed parts, material wastage, and inefficiencies in production processes due to some problems that will be mentioned afterward. This project aims to address these issues by developing a machine-learning-enhanced digital twin concept for real-time defect detection and prediction. The proposed digital twin concept integrates machine learning models, simulations from ANSYS Workbench, and a digital twin of the DED printer into one system that will predict the crack anomalies in an object. The machine learning models are trained on simulated data that will be taken from ANSYS to predict defects like cracks and deformation based on input parameters. A user-friendly software interface will connect all components of the system. It enables real-time monitoring and control over the printing object and lets us choose optimized input parameters that will give no cracks. Future work will include designing a detailed 3D model of the DED printer in SOLIDWORKS and expanding the database with experimental data collected from real-world printing.

Moreover, the machine-learning algorithms will be refined for improved accuracy. By validating the system through physical tests, this project seeks to enhance the reliability, efficiency, and sustainability of the DED 3D printing process. That will contribute to its wider adoption in industrial applications.

Acknowledgment

We would sincerely appreciate the assistance of all those who contributed to the successful finalization of this thesis. Our deepest gratitude goes to our supervisor of the project, Professor Essam Shehab, and co-supervisor, Professor Md. Hazrat Ali, for their immense guidance, constant support, and constructive criticism during this whole project. Their inputs and encouragement have been a beacon to our work. There is no way we can express our extreme gratitude to Nazarbayev University for giving us the opportunity of thesis work, and the necessary resources and facilities for conducting our research. The world-class education and the supportive academic community have been the beds of our growth. We will forever remain grateful to all the professors, instructors, and supporting staff members who have advised and nurtured us throughout our academic journey. Their dedication has made an impact not only on this project but also on our general intellectual development. Last but not the least, we will forever be indebted to our families for their love, patience, and encouragement. Their belief in us when the going got tough has been our greatest strength and motivation.

Table of Contents

Chapter 1. Introduction

1.1 Background

Additive manufacturing or AM has rapidly become one of the most innovative approaches for shaping complex geometries, which enable part consolidation, design flexibility, and on-demand production. One of the AM techniques, Directed Energy Deposition (DED), has drawn the most attention due to its capability to process high performance metals and restore/rebuild existing components. DED not only allows for near-net-shape fabrications with customized material properties by fusing material as it is being deposited using a concentrated energy source through a laser or electron beam, but it also can use valuable alloys such as Inconel 718 to fabricate aerospace, and tooling components. However, there are several issues of the DED printers including thermal stress buildup, residual deformation, material wastage, and especially defect formation, such as cracking, oxidation, porosity, and warping, which restrict the wider industrial adoption. Complex thermomechanical interactions that take place during rapid thermal cycling, especially in multi-layered builds, are the driving force behind these phenomena. The combination of machine learning (ML) and digital twin frameworks has become a cutting-edge strategy to overcome these challenges. Real-time simulation, monitoring, and prediction of manufacturing process behavior are possible with digital twins, which is a virtual representation of the physical system. Digital twin (DT) can help to improve part quality and process reliability by anticipating process anomalies and proactively optimizing input parameters when enhanced with machine learning algorithms trained on simulations or sensor data. A machine learning framework enhanced by digital twins is created in this project to identify and forecast DED flaws. The Meltio M450 (Figure 1), is a small, industrial-grade wire-laser DED system, that serves as the basis for the simulation and validation activities in this project. Utilizing wire and powder feedstock, the Meltio M450's multi-laser diode system can process a variety of metals, including Inconel 718, with high deposition efficiency and ease of material handling. Its total output capacity can reach values up to 1200 W. The M450 functions as the digital reference for simulation and DT development as well as the experimental platform due to its compatibility with G-code, precise motion control, and integrated feedback sensors. This project intends to aid in the creation of intelligent and dependable DED systems for next-generation manufacturing applications by fusing machine learning, experimental hardware, and numerical simulation under a single digital twin concept.

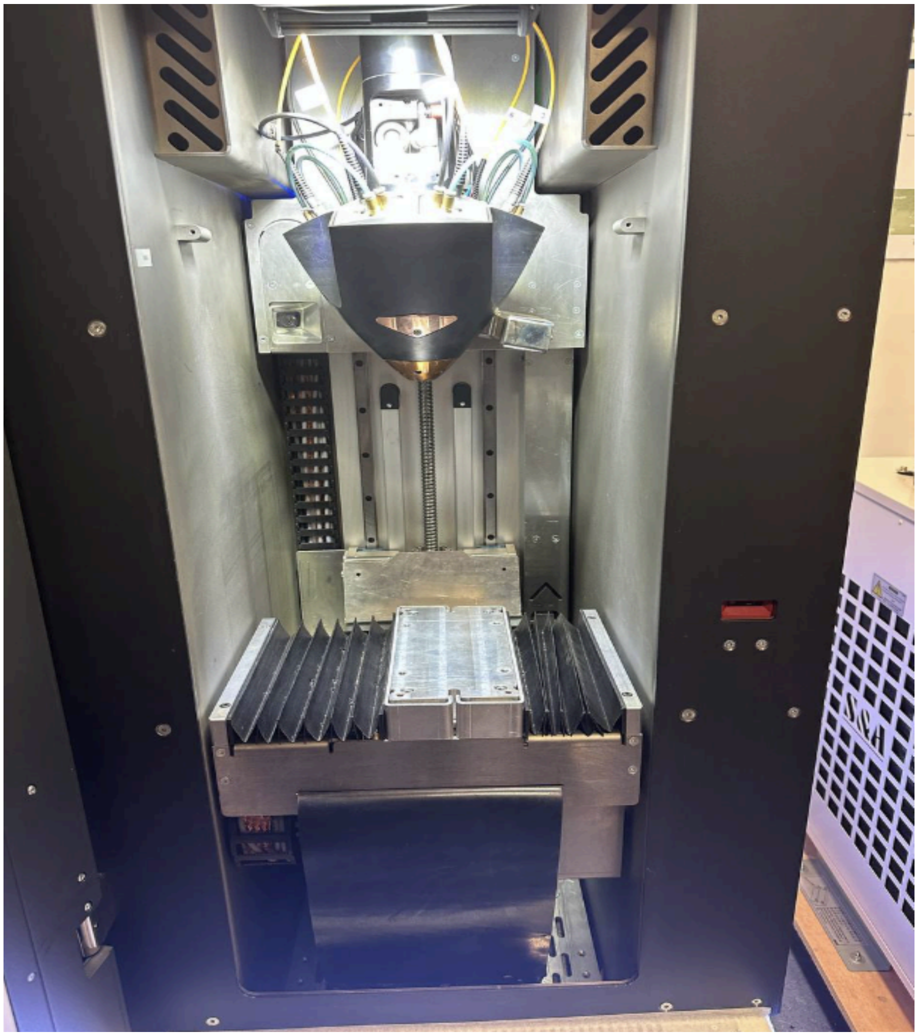


Figure 1. Directed Energy Deposition (DED) printing machine.

1.2 Problem Statement

Despite the fact that 3D metal printing, especially direct energy deposition printing technology has promising potential to create intricate geometries and can be used in order to repair expensive parts within the small systems, It has considerable obstacles that are mentioned below.

1. **Defects in Printed Components**: Small deformations, fractures, cracks and other defects are structural flaws that weaken or even destroy the mechanical integrity of the parts, not

only reduces the quality of the printed parts, but also may lead to the catastrophic failures during the operation.

2. **Material and Energy Wastage**: Defected parts can not be used further, and used material can not be recycled without external expenses, which cause material and energy wastage increasing not only the cost of the part, but technology of printing itself.
3. **Lack of Defect Detection and Prediction**: The fact that there are no methods to detect or predict the defects in metal 3D printing, any small variables could cause the entire part to fail.

Figure 2 demonstrates the defect formation in a variety of cases, starting from oxidation and ending up with crack formations. All of the defects mentioned above may cause system failures, or the printed part will not perform as it was designed for. It reduces the efficiency and the cost of this technology of metal 3D printing.

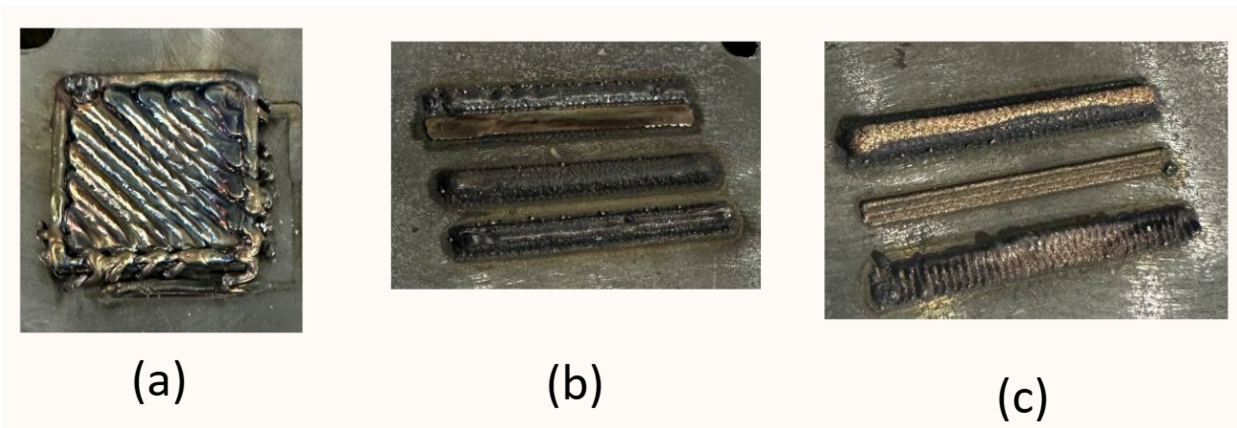


Figure 2. Defects in Metal 3D Printing (a. Oxidation, b. Lack of Fusion, c. Crack Formation)

These obstacles restrict the application of metal 3D printing in industry and reduce its competitiveness compared to other alternatives, such as using full overhaul of the system or particularly on devices. In order to deal with these issues, defect prediction and real time detection algorithms are needed.

Real time monitoring, predicting algorithms on the basis of optimization of input parameters could be integrated through the Machine Learning (ML) algorithms. All of these measures could decrease the material wastage and defect formation in printed parts.

1.3 Aim and Objectives

The main aim of this research project is to create a machine-learning based digital twin conceptual model for metal direct energy deposition 3D printing, which will stand for real-time failure prediction and detection.

The main objectives are to:

1. Develop an accurate machine learning model to assess the risk of defects and estimate overall deformation in printed components.
2. To anticipate and avoid any problems during the printing process, use predictive analytics.
3. Design and implement a digital twin system concept for the DED printer to monitor the printing process in real-time.

1.4 Report Structure

The main body of the report is organized into several chapters. It starts with an introduction that consists of the overall background of the project, existing problems, aims and objectives of the project, and briefly summarizes the advantages and disadvantages of the project. Following part is a review of the related studies in this field and the usage of the historical data from those researches. The Methodology part introduces the brief description of the whole process, including data collection, building machine learning models, and construction of the digital twin concept. It is followed up with a results part, that demonstrates the outcome of the digital twin concept, including all of its limitations and how accurate it is, explaining how some specific parameters impact the defect formation. Conclusion part summarizes the project, briefly mentioning the achievements and states future work that could be done in order to improve the system and the project itself. Appendix and References parts consist of all links to the sources and almost every code, and other related figures that are not placed in the main parts of the report.

2. Literature review

Lalegani et al. (2022) reviewed additive/subtractive hybrid manufacturing of directed energy deposition (DED) processes with different metal powder and wire configurations and proposed their varying results. Accuracy and precision of the machine was lacking and therefore authors ended up proposing a new method which is called ASHM. That method utilizes a multi axis nozzle deposition process. With this new method in use the porosity decreased by 81% which is a significant improvement from the previous method.

"Processing Challenges and Delamination Prevention Methods in Titanium-Steel DED 3D Printing" published by Andreu et al. shared valuable information about challenges in DED printers. It is very common for the crack to occur in titanium alloys as they are very brittle in their nature and thus material selection is of high importance. Also, preheating the substrate up to ~800K is very good in enhancing the titanium aluminide.

Raghavan and Duda (2016) published the 3D Metal Printing Technology which gives an overview of all types of 3D printing techniques such as powder bed fusion, laser deposition method and others. It shows all the main differences in each type of method such as preheating occurrence or material selection.

Svetlizky et al. published "Laser-based directed energy deposition (DED-LB) of advanced material". It discusses the importance of the laser based printing technologies and their market increase in the last decade. Main things discussed are material design and a variety of material compositions.

Choi et al. "A Study on Strengthening Mechanical Properties of a Punch Mold for Cutting by Using an HWS Powder Material and a DED Semi-AM Method of Metal 3D Printing" that discusses strengthening the mechanical properties of the model. They used high wear resistance steel. Both the toughness and wear resistance were 4 and 4.2 times better than D2 specimens that are common in the DED printing field.

Zaubier and Radi published “Finite Element Modeling (FEM) of the thermal behavior of 3D printed parts during Directed Energy Deposition (DED)” that uses ANSYS Workbench to create the model. The numerical model in the Workbench predicts several properties such as temperature, pool dynamics and others.

Anne published (2022) “Modeling of the Selective Laser Melting (SLM) Process Using the ANSYS Software” which uses laser based 3D printing methods in the Ansys Workbench. Ansys' Additive Science addition was used specifically to create the model and to see the all advantages and disadvantages of the SLM method.

“Three-Dimensional Numerical Simulation of Laser Metal Deposition Method with Bead Profile Optimization” was published by Maurya et al. in 2024. They created a model in ANSYS using a high deposition rate to see if there will be any crack or some impact on the model itself.

Chapter 2. Literature review

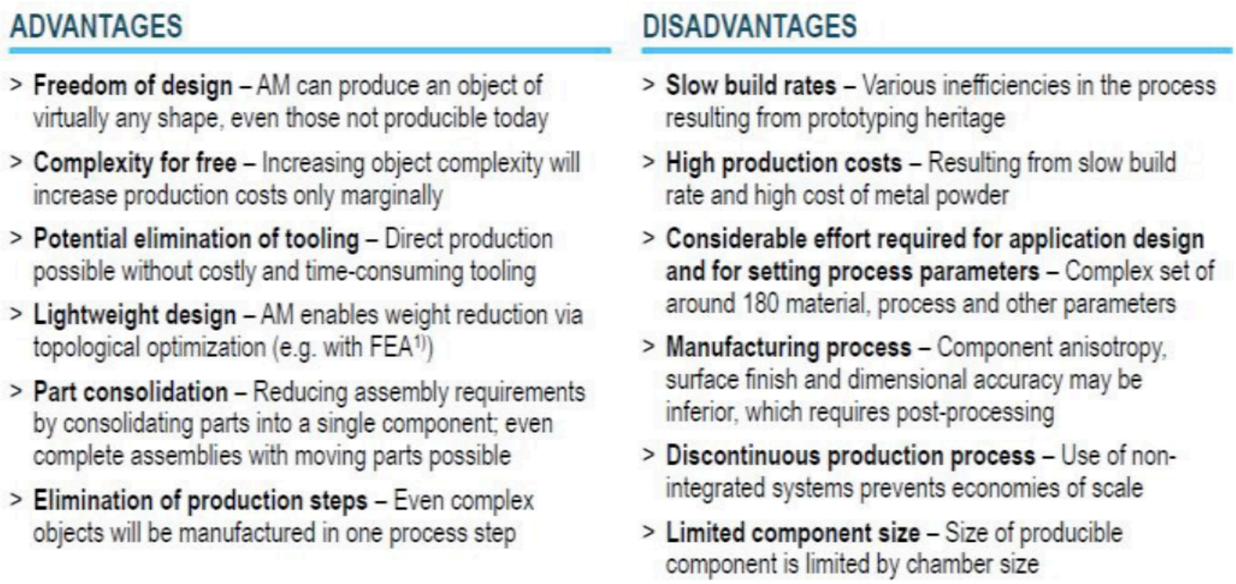
2.1 Introduction

- General overview of Additive Manufacturing (AM) in metals.
- Importance of defect prediction, simulation, and digital twin technologies.
- State why this field needs integration of physics + ML + twin systems.

2.2 Additive Manufacturing Technologies

- 2.2.1 Metal Additive Manufacturing Processes

In the 3D printing world, Additive Manufacturing (AM) defines the technique, where a digital model is sliced and printed layer by layer. The benefits of this technology are light weighting, part consolidation, high precision and rapid prototyping [Duda & Raghavan, 2016]. Metal additive manufacturing is found to be used in aerospace, marine engineering, and conventional automotive industry, where metal needs to be in a complex, accurate shape. Meanwhile, the limitations include component size, high production cost, and need for post-processing. The quick list of advantages and disadvantages can be found in figure X, represented by Duda & Raghavan in 2016:



ADVANTAGES

> **Freedom of design** – AM can produce an object of virtually any shape, even those not producible today

> **Complexity for free** – Increasing object complexity will increase production costs only marginally

> **Potential elimination of tooling** – Direct production possible without costly and time-consuming tooling

> **Lightweight design** – AM enables weight reduction via topological optimization (e.g. with FEA[1)])

> **Part consolidation** – Reducing assembly requirements by consolidating parts into a single component; even complete assemblies with moving parts possible

> **Elimination of production steps** – Even complex objects will be manufactured in one process step

DISADVANTAGES

> **Slow build rates** – Various inefficiencies in the process resulting from prototyping heritage

> **High production costs** – Resulting from slow build rate and high cost of metal powder

> **Considerable effort required for application design and for setting process parameters** – Complex set of around 180 material, process and other parameters

> **Manufacturing process** – Component anisotropy, surface finish and dimensional accuracy may be inferior, which requires post-processing

> **Discontinuous production process** – Use of non-integrated systems prevents economies of scale

> **Limited component size** – Size of producible component is limited by chamber size

Figure X1. Advantages and Disadvantages comparison

There are two main classifications for Metal AM - Powder Bed Fusion (PBF) and Direct Energy Deposition (DED). If PBF uses thermal energy to fuse the region of powder bed, DED uses thermal energy to melt material at the moment of deposition. In practical use, DED is capable of printing bigger part sizes rather than PBF method, and best for repair and modification of components in production.

- 2.2.2 Directed Energy Deposition (DED) Process

To explain deeper the method of DED, its thermal energy is usually taken from laser, electron beam or plasma arc [Andreu et al., 2024]. It has further sub-classifications regarding how thermal energy is derived, but the general idea is that this energy is used to melt material at the depositing moment. Here, the material might be a typical metal powder input to the focal point of energy through a coaxial nozzle, or might be delivered with a wire feeder to the exact point on the substrate. Worth to note, that inert gas delivery is also compulsory, as it prevents just melted metal from oxidation. The nozzle follows a pre-calculated path during printing, as deposited metal cools down and solidifies gradually. The schematic work process of DED can be observed in Figure X provided by Andreu et al. (2024).

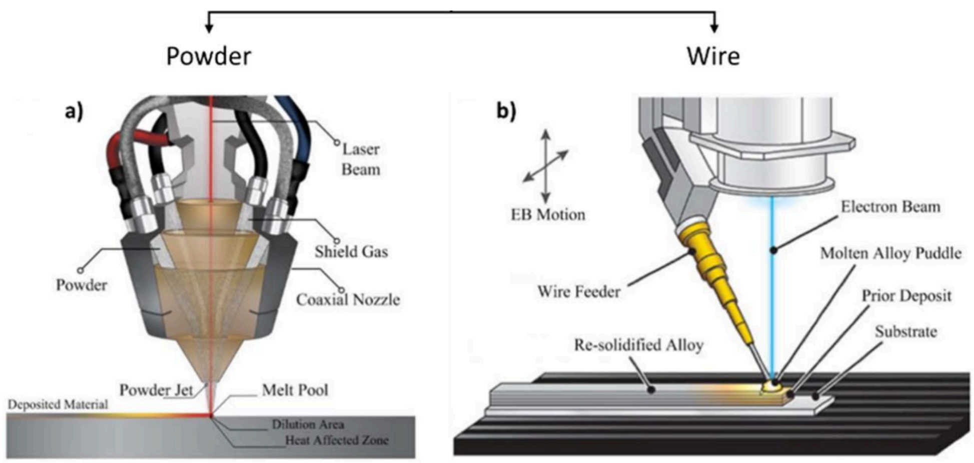


Figure X2. Schematics of a) Powder based and b) Wire feeder DEDs

- 2.2.3 Common Defects in DED Printing

Apart from the disadvantages demonstrated in Figure X1, the DED has common defects within printing processes. Understanding and studying the common defects is crucial in the way of improving the overall quality of 3D printers. When metal is melted, then solidified, complex chemical and thermal reactions happen, each of which, in case of poor calculation, can cause the failure of printed parts. For instance, porosity oftenly happens due to inhomogeneities in printed parts [Dezaki et al., 2022]. Any unexpected change in microstructure of material causes defects. Another major problem is delamination, which occurs when deposited layers do not solidify into a consistent object, with further outcome of crack or separation of layers [Andreu et al., 2023]. Oxidation might also occur, when the delivery of inert gas is reduced or damaged. The diagram in Figure X3, presented by Dezaky et al. (2022), shows that the most common failure in DED printing is Porosity, Oxidation and Crack. Additional minor defects are classified by DED printer types.

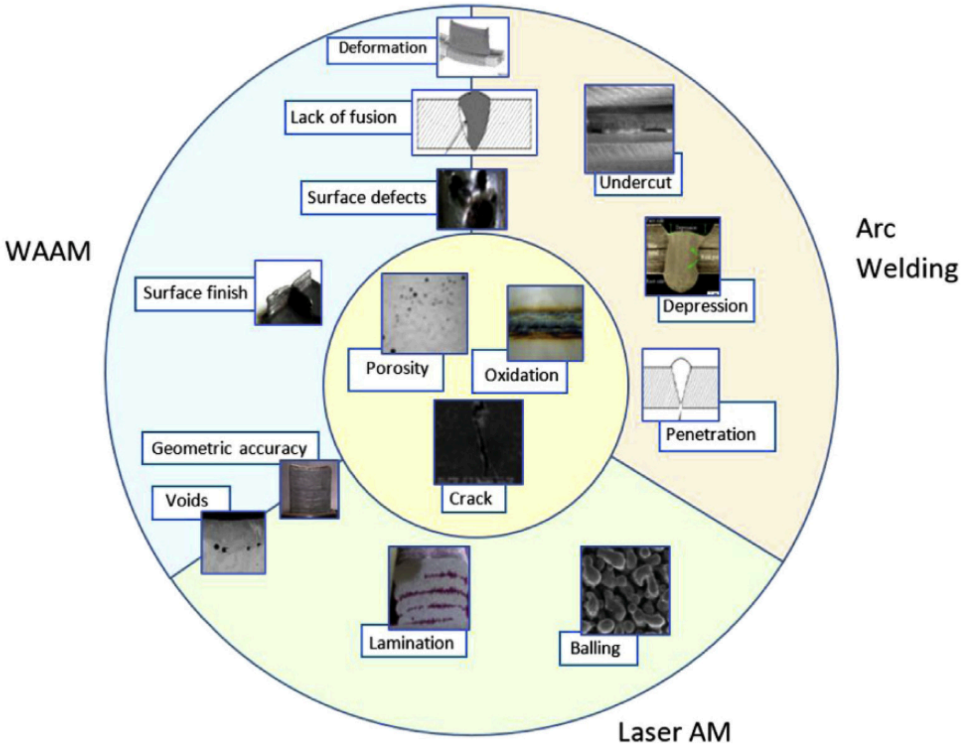


Figure X3. Common Defects in DED Printing

- 2.2.4 Process Parameters Influencing Defects

Meanwhile, the vast majority of defects can be studied, and mainly caused due to improper input parameters which we mostly control. Andreu et al. (2023), in their study concluded that preheating of substrate can effectively help to reduce the residual stresses. There is a link between preheat temperature and metal solidification behavior, that if the comfortable temperature is set on substrate before printing metal, the delamination can be avoided. Figure X4 also demonstrates that preheating has effect on crack occurrence at microstructure level:

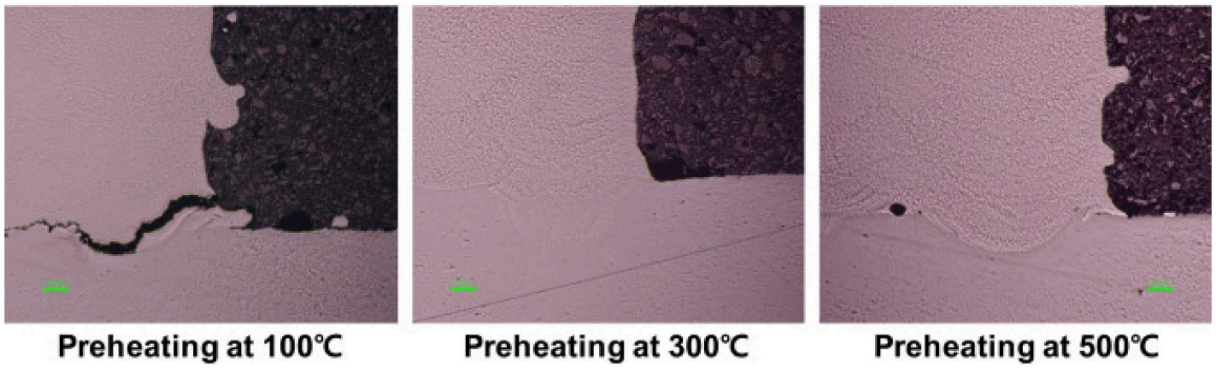


Figure X4. Prehearing and microstructure of bounds

Nevertheless, the surface defects have a strong link with the feed rate. Dezaki et al. (2022), suggested that capability of controlling the feed rate leads to better surface quality. Moreover, they provided how Laser Power and Scanning Speed have effect on metal binding, which can be observed in Figure X5:

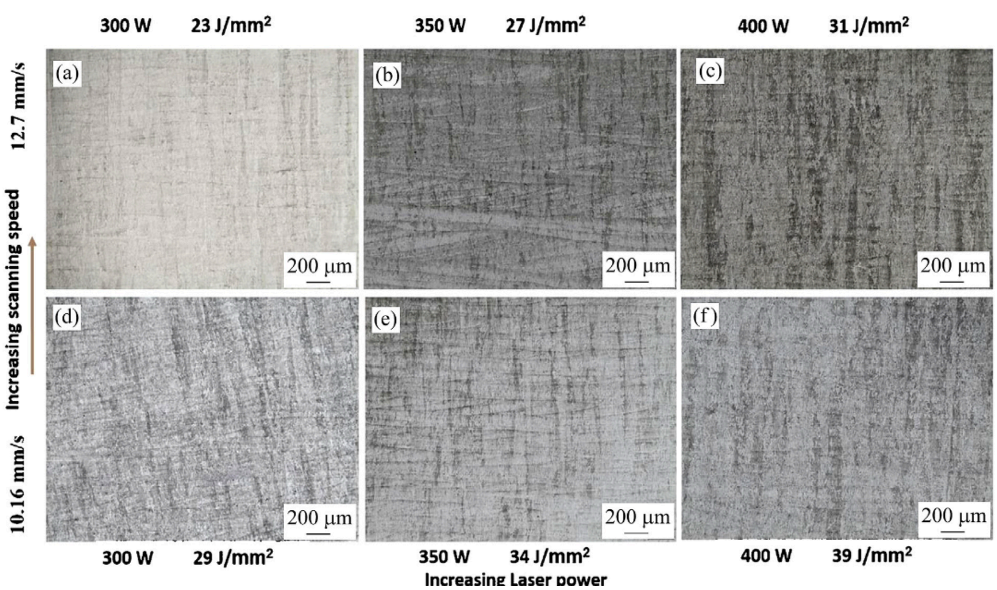


Figure X5. Laser Power and Scanning Speed relationship

These insights can help us to successfully predict and avoid most deformations to occur in metal 3D printing, that defects depend on input parameters, and in our case they are RPM (Scanning Speed), Feed Rate and Laser Power.

2.3 Thermomechanical Simulation in Additive Manufacturing

- 2.3.1 FEM Modeling Techniques
    - Thermal–mechanical coupled simulations.
- 2.3.2 Stress and Deformation Prediction
    - How thermal gradients drive mechanical defects.
- 2.3.3 Simulation Challenges
    - Computation time, mesh quality, real-time limitations.

✅ Focused discussion:
Summarize and compare different simulation approaches, highlight your ANSYS choice.

2.4 Machine Learning for Defect Prediction

- 2.4.1 Supervised Learning Applications

Supervised learning has become an important tool in metal additive manufacturing, in particular for defect detection and process optimization. Supervised learning models predict possible defects, given a combination of such parameter settings as laser power, scan speed, and material feed rate, which have been previously labeled with defect outcomes. These models unravel past process data and find patterns leading to defects such as cracks, porosity, or warping. They can therefore, once trained, investigate combinations of new parameters in real time so they can inform decisions to improve print quality.

Regression-based predictions of defect probabilities are the most common methods in this field of application for supervised learning. Algorithms like linear regression, support vector regression (SVR), and neural networks are trained using simulation or experimental data to predict the likely defects using process variables as input. For instance, the model can predict that a certain setup of high laser power and low scan speed gives rise to an increased likelihood of keyholing or excessive residual stresses, encouraging the manufacturer to adjust parameters beforehand, thus reducing material wastage and increasing reliability of the components.

The second key application for classification models provides a mechanism for supervised learning to designate printed parts as either "defective" or "non-defective" based on sensor signals. Such methods as logistic regression, random forests, or convolutional neural networks (CNNs) can analyze thermal imaging and acoustic emissions or direct layer-wise visual information to flag anomalies while printing. This dynamic monitoring allows for on-the-go correction, thereby reducing demands for post-printing inspection and reconditioning.

Another area where supervised learning contributes is parameter optimization. Examining successful print data offers a guide for models to suggest settings that will work best for new geometries or materials, consequently cutting down trial-and-error. For example, algorithm-wise gradient boosting machines (GBMs) could be used to model non-linear dependencies between parameters and outcomes and provide recommendations for modifications that optimally trade-off speed, energy consumption, and strength of parts.

In the world of metal 3D printing, supervised learning is greatly preferred for its accuracy and flexibility. Unlike unsupervised learning, which searches for hidden patterns lying in unlabeled data, supervised processes draw predictions from unique examples of known defects in a prompt and useful manner. This gives them a key position in quality control in an industry where defects can lead to costly failures. As datasets and algorithms evolve through gradual improvement, the future of supervised learning shall pave the way for reliable and high-quality additive manufacturing.

Key Advantages Over Unsupervised Learning:

Compared to unsupervised approaches, supervised learning gives more direct and actionable recommendations for applications in metal 3D printing. The primary difference is in the treatment of the data and the types of results that are derived from them.

With supervised learning, we prepare models using labeled data for which we already know the process parameters that produced good or bad prints. It enables specific, quantifiable predictions, like, "there are 87% chances of porosity if these settings are used." It also gives specific suggestions about parameter changes to use in order to avoid defects and can even identify which parameters are most likely to be talent scouts for a particular problem. We can easily measure how accurate these models are with a metric such as R^2 scores and root mean square error (RMSE).

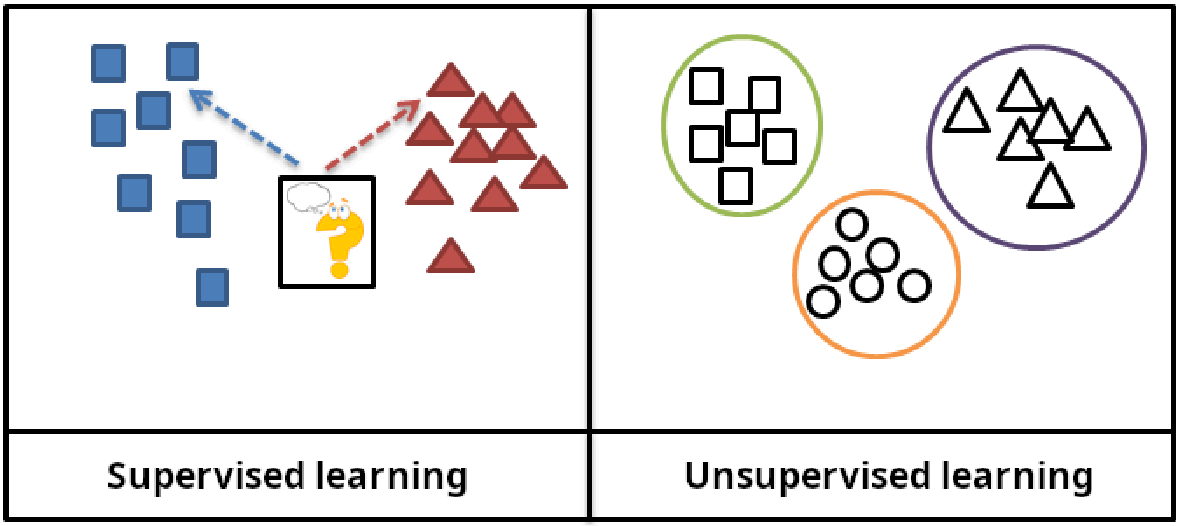


Figure 3. Basic illustration of the difference between two learning techniques

Unsupervised learning works with latent patterns in the collected data, from which no labels exist. This would be very effective in initial exploration of the data, but with some limitations for example, it would group printing runs that appear similar but could not tell which ones belong to the group of defective parts unless someone puts some manual inspection.

It also makes supervised learning especially significant for precision-critical applications in manufacturing. In industries like aerospace or medical devices, one needs to know exactly what causes defects and how to eliminate them-not simply that certain patterns exist. Such supervised models would enable direct linking of specific temperature fluctuations or laser settings to particular types of flaws, thereby creating more reliable quality control.

- 2.4.2 Linear regression, Ridge regression

Establishment of a linear regression relationship among process parameters and defect-inducing quality metrics is done by the equation as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon$$

whereby Y acts as defect metrics (porosity percentage, surface roughness, residual stress), X_i are process parameters (laser power, scan speed, layer thickness) whereas β_i become coefficients quantifying the effect of parameters. With this simple yet very powerful formulation, it becomes possible to interpret how each parameter has an effect on the final quality of the parts.

In production environments these Linear regression techniques represent advantages from the operational perspective:

1. Gives clear parameter weighting
2. Coefficient directly interpretable with respect to physical concepts
3. Excellent baseline for development of processes
4. Computationally very efficient for real-time applications in control

Ridge regression modifies this structure into L2 regularized one:

$$\hat{\beta} = \operatorname{argmin}(\|Y - X\beta\|^2 + \lambda\|\beta\|^2).$$

The penalty term $\lambda\|\beta\|^2$ defines several issues concerning metal AM's future.

1. Multicollinearity among some interdependent process parameters
2. High-dimensional parameter spaces with few experimental data
3. Numerical instability caused with matrix inversion operations

Ridge regression offers one of the most impactful and beneficial advantages:

1. Keeps interpretation intact for the model but increases robustness.
2. Magic balancing of influences of parameters.
3. Less sensitive to measurement noise.
4. Careful λ tuning needed by cross-validation.

Most commonly, ridge regression excels in online production and quality assurance systems as its stability holds significant account in monitoring builds across several machines or natural variations affecting the properties of materials. Most of the production systems in use actually combine both; linear regression for interpretable process insights during development, then ridge regression for robust quality control during volumetric production.

- 2.4.2 Data Augmentation in AM (GANs)

Definitely one barrier to developing strong machine learning models is the lack of experimental data in metal additive manufacturing research. Thus, Generative Adversarial Networks (GANs) become indispensable, especially for very expensive and time-consuming processes like Directed Energy Deposition (DED), Powder Bed Fusion (PBF) and etc.

The Emergence of a Need for GANs in Our Research Into AM for Metals

In our research on metal additive manufacturing, we face difficulties with data, which strongly qualifies GANs as a tool. Experimental work causes basic problems that are hard to overcome using conventional data collection, thus coverage for synthetic data generation needs to be better.

Physical and Financial Barriers

Metal additive manufacturing processes are typically prohibitively expensive and time-consuming to investigate experimentally. Every single build consumes hundreds of dollars of specialized metal powders and significant post-processing and characterization. Machine time alone costs over $150 per hour for industrial systems, making a comprehensive parameter study prohibitively expensive. The best GANs can do is constructively in between that by delivering high-quality synthetic data while expanding the limited experimental dataset while keeping the fundamental physical relations between process parameters and outcomes.

Scarcity of Data and Its Imbalance

Our existing physical experiments dataset is both small and imbalanced: we have many examples of successful builds, whereas failure modes and defect cases are significantly under-represented. This imbalance challenges building training sets for accurate machine learning models since most of them would require success in many cases to get a few failures. GANs address this by intelligently synthesizing these rare but important cases, which would otherwise require hundreds of additional physical trials to capture adequately.

Testing High-parameter Spaces with Security

The high-dimensional nature of metal AM processes with many interacting variables leads to an overhead in parameter space that cannot be completely tracked out in physical experimentation. The important area for physical testing can thus be identified, and costly and dangerous parameter combinations that would destroy equipment or waste materials are avoided. This accelerates our research effort and cuts down costs and risks.

Better Model Development and Generalization Strength

By using GANs to create a repertoire of diverse, physics plausible training examples, we are helping our machine learning models to build up a more robust understanding of the underlying process physics. This would result in better generalization performance when exposed to new materials, geometries, or machine configurations. The synthetic data acts like "virtual experience," supplementing our limited physical experiments.

Challenge	Traditional Solution	GAN Solution	Benefits
Limited experimental data (34 builds)	Reduce model complexity	Generate 1000+ synthetic samples	6x more training data
Missing defect examples	Oversample rare cases	Create physically-plausible defect variants	Better defect detection accuracy
High-dimensional parameter space	Constrain variables	Explore full parameter space virtually	Discover optimal regions safely
Costly trial-and-error	Sequential DOE	Simulated parameter screening	70% fewer physical trials needed
Noisy sensor data	Manual cleaning	Learn noise patterns for robust generation	More reliable synthetic data

Table 2. GAN Solutions vs Traditional Approaches

Table below represents the main sources and their key features in ML implementation in our project.

Table 3. ML Literature Review

№	Source	Key findings
1	Jin, L., Zhai, X., Wang, K., Zhang, K., Wu, D., Nazir, A., Jiang, J., & Liao, W. (2024). Big data, machine learning, and	Advantages: Comprehensive overview, detailed application of ML, digital twin

	digital twin assisted additive manufacturing: A review. *Materials & Design, 244*, 113086. https://doi.org/10.1016/j.matdes.2024.113086	in metal AM. Disadvantages: Lacks Practical implementation details, and limited focus on real-time feedback loops.
2	DebRoy, T., Mukherjee, T., Wei, H. L., Elmer, J. W., & Milewski, J. O. (2020). Metallurgy, mechanistic models, and machine learning in metal printing. *Nature Reviews Materials, 6*(1), 48–68. https://doi.org/10.1038/s41578-020-00236-1	Advantages: Detailed Focus on Metal Additive Manufacturing, Emphasis on Machine Learning for Defect Detection, Insight into Metallurgical Factors Disadvantages: Lacks Practical Case Studies, Limited Discussion on Digital Twin Technology
3	Min, Q., Lu, Y., Liu, Z., Su, C., & Wang, B. (2019). Machine learning-based digital twin framework for production optimization in the petrochemical industry. *International Journal of Information Management, 49*, 502–519. https://doi.org/10.1016/j.ijinfomgt.2019.05.020	An effective solution to address series data processing issues, results show the effectiveness
4	Su, X., Yan, X., & Tsai, C.-L. (2012). Linear regression. *Wiley Interdisciplinary Reviews: Computational Statistics, 4*(3), 275-294. https://doi.org/10.1002/wics.1198	Linear regression basics and data will be used in the process of creating an ML algorithm.
5	Zeng, L., Wu, W., & Li, Z. (2023). Research on multiple regression linear prediction model of 3D printing process parameters. *In X. Yuan & G. Wu (Eds.), Third International Conference on Mechanical, Electronics, and Electrical and Automation Control (METMS 2023) (Vol. 12722, p. 127224Q). SPIE.* https://doi.org/10.1117/12.2679759	Multiple Regression Mathematical Model, This model can be interpreted in an ML algorithm that could be used.
6	Khan, M. F., Alam, A., Siddiqui, M. A., Alam, M. S., Rafat, Y., Salik, N., & Al-Saidan, I. (2021). Real-time defect detection in 3D printing using machine learning. *Materials Today: Proceedings, 42*, 521-528. https://doi.org/10.1016/j.matpr.2020.10.482	CNN Model is used here. We are not sure about this source because we have not decided whether to use sensor data or images and deep learning.
6	Goh, G. D., & Yeong, W. Y. (2021). Applications of machine learning in 3D printing. *Materials Today: Proceedings, 42*, 521-528. https://doi.org/10.1016/j.matpr.2020.10.482	Applications of machine learning in 3D printing examples
7	Sampedro, G. a. R., Putra, M. a. P., & Abisado, M. (2023). 3D-AmplifAI: An ensemble machine learning approach to digital twin fault monitoring for additive manufacturing in smart factories. IEEE Access, 11, 64128–64140. https://doi.org/10.1109/access.2023.3289536	The 3D-AmplifAI algorithm is a key strength, as it increases prediction accuracy by leveraging the strengths of multiple ML models. Although effective for fault detection in FDM printers, the study doesn't extensively

		explore applications in metal 3D printing

2.5 Digital Twin Technology for Additive Manufacturing

- 2.5.1 Definition and Components

  Digital Twin and Additive Manufacturing (AM) are both key technologies of the fourth industrial revolution. Primary purpose of additive manufacturing is to produce complex components or products that are difficult to manufacture using conventional methods. However the quality of the printed products is still a significant issue. Digital twin is one way of solving these problems through real-time monitoring, controlling of the processes.
- 2.5.2 Applications in Metal AM

  There are several applications of digital twin implementation in additive manufacturing. Firstly, DT implementation can be used in order to monitor the system. It connects directly to the system and represents the values from sensors, cameras and other sources of information that system might have. Secondly, It can be used in order to predict defect formations through the analysis of received data. It could be both analysing already received data from previous works or real-time monitoring. In addition to applications mentioned above, this technology in Digital Twin could be used in adaptive controlling of the system. Dynamic adjusting of input parameters could improve the quality of the products.
- 2.5.3 Challenges

  There are lists of the challenges related with DT integration in additive manufacturing processes. They depend on the systems, printers that they are using for, but genuinely there are three main points stated below. Sensor Integration or usage already in-build sensors and cameras. First of all, Digital Twin is the data link between the real and digital systems. The only way to receive real-time information is to use either in-built or integrated sensors to transfer data about the key parameters. Besides the sensor part there are challenges related to the synchronization. Not all systems could be connected directly and real-time data synchronization could be a big issue for systems with closed software. Last but not least is the computational power of the hardware. Increase in the system complexity requires more computational power of the device.

Following table represents the source that we used.

Table 4. DT Literature review.

№	Source	Key findings
1	Mukherjee, T., & DebRoy, T. (2019). A digital twin for rapid qualification of 3D printed metallic components. *Applied Materials Today, 14*, 59–65. https://doi.org/10.1016/j.apmt.2018.11.003	Real-Time Monitoring, Defect Prediction, Machine Learning Integration, Major metallurgical challenges, Data-Driven Errors
2	Jyeniskhan, N., Keutayeva, A., Kazbek, G., Ali, M. H., & Shehab, E. (2023). Integrating machine learning model and digital twin system for additive manufacturing. *IEEE Access, 11*, 71113–71126. https://doi.org/10.1109/access.2023.3294486	Advantages: Comprehensive Integration Framework, High Detection Accuracy, Real-Time monitoring and control. Disadvantages: Limited to FDM Printers, Lack of Standardization
3	Malik, A. W., Mahmood, M. A., & Liou, F. (2024). Digital twin–driven optimization of laser powder bed fusion processes: a focus on lack-of-fusion defects. Rapid Prototyping Journal. https://doi.org/10.1108/rpj-02-2024-0091	Advantages: Predictive Modeling and Control, Innovative Digital Twin Architecture Disadvantages: Limited Material and Process Scope (AISI 316L only)
4	Kantaros, A., Piromalis, D., Tsaramirsis, G., Papageorgas, P., & Tamimi, H. (2021). 3D printing and implementation of digital twins: Current trends and limitations. Applied System Innovation, 5(1), 7. https://doi.org/10.3390/asi5010007	A comprehensive review of various DT applications in metal 3D printing, and also discusses integrating computational models and sensor data to enhance the accuracy and predictive capability of DTs.

2.6 Research Gaps Identified

Our thorough documentary research discovered a number of critical knowledge gaps in contemporary research about Directed Energy Deposition (DED) printing. This gap brings us to address the capstone project. Although considerable work is ongoing in isolated components of DED processes, studies integrating most, if not all, components into a working, functional system remain rare.

- Poor Process Control Implementation

Present DED printing research highlights the disconnect between defect prediction capability and the real-world implementation of closed-loop process control. Many studies establish impressive machine learning models for defect detection based on high accuracy; however, these developments are still not fully operational, more theoretical instead. Most published papers follow the detection paradigm only, warning or grading the part but not integrating the defect control into any form of feedback mechanism for real-time prevention. This is despite the strong industrial need for capabilities that will enable autonomous process adjustments.

Current approaches demonstrated by the literature has three fundamental shortcomings. First, an significant implementation gap exists, where academic models hardly ever proceed beyond laboratory demonstrations. In our review, we found only two situations in which a feedback loop between defect detection and process correction was established, both in very limited experimental setups. Secondly, the majority of systems do not contain interfaces for changing key printing parameters such as laser power or scan speed dynamically during printing. Thirdly, the existing solutions do not address the real-world challenges of industrial deployment, including sensor integration, computational latency, and reliability requirements.

- Most studies focus on one or two parts (simulation only, or ML only), but few combine all three.

One of the most obvious gaps is the complement missing between simulation, machine learning, and live digital twin implementation. Most of all published research has focused on only this aspect: either creating high fidelity thermal simulations, developing machine-learning independent models for defect prediction, or suggesting conceptual digital twin frameworks.

These disconnected approaches fail to capture the full potential of combining physics-based modeling and data-driven analytics within an operational production environment. For instance, while many studies portray the melt pool dynamic correctly in ANSYS, very few take this data into real-time predictive models-and an even smaller fraction connects these predictions to operational digital twin interfaces.

- Very little real-world Unity or real-time visualization in DED systems.

Chapter 3. Methodology

This project adopts an integral methodology: advanced simulation, machine learning, hardware design, and validation chain in developing a robust system for defect detection and prediction in metal 3D printing. High-fidelity simulations are conducted using the ANSYS Workbench in replicating the DED process to generate critical datasets on thermal and mechanical behaviors. Beyond this, such datasets will now constitute the backbone of the machine learning model that is supposed to support prediction and mitigation regarding defect appearance, considering some key process parameters.

The subsequent stage will consist of a digital twin framework development: integrating the 3D model of the printer, real-time data input, and visualization and control by means of Unity. Such integration will make the machine learning model, the digital twin, and the physical printer all communicate seamlessly. The methodology will be concluded by rigorous testing and validation, where the performance of the system will be assessed in simulated and real-world conditions to ensure its accuracy, reliability, and scalability for wider applications in additive manufacturing.

3.1 Ansys Simulations

A key component of the suggested defect detection and prediction system for metal additive manufacturing is the simulation workflow. That is why the creation of synthetic data that faithfully captures the physical behavior of the Directed Energy Deposition (DED) processes under controlled circumstances was one of the main goals of this project. ANSYS 2024 R1 Workbench was used to perform high fidelity simulations that included coupled thermal/mechanical models to model printing-related temperature gradients, stress accumulation, and deformation behavior. Previous studies by Zaubier and Radi [18] and Mukherjee and DebRoy [1] employed similar simulation techniques, supporting the methodology of this project.

The process begins with defining key input parameters such as laser power, material deposition rate, and ambient temperature. Defects such as cracking, oxidation, and layer misalignment are closely related to residual stresses and thermal gradients, which are greatly influenced by these factors [4], [14].

As illustrated in Figure 3, the simulation pipeline includes the following five stages:

1. Definition of input parameters
2. Geometry creation and meshing
3. ANSYS simulation of DED thermal and mechanical behaviors
4. Structured data export and filtering
5. Normalization and validation for machine learning readiness

The outputs of the simulation include variables such as temperature distribution (°C), von Mises stress (MPa), shear stress (MPa), and total deformation (mm), which were identified as most impactful indicators to defect formation based on metallurgical studies by DebRoy et al. [4] and Andreu et al. [14]. In addition, these results were exported in structured CSV and XLSX formats and processed using statistical filtering and feature scaling, enabling compatibility with the machine learning framework developed in later stages. In line with current developments in simulation-based machine learning for additive manufacturing, the process not only accurately simulates real-world conditions but also generates a reliable training dataset for predictive modeling [2], [3], and [20].

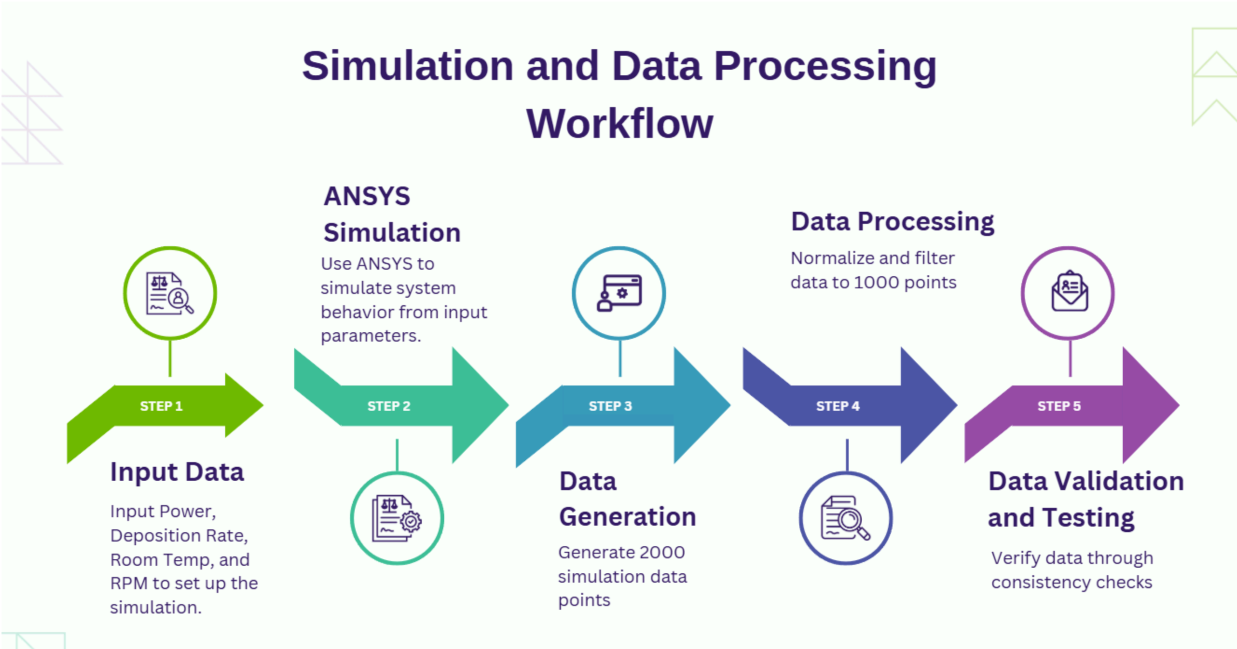


Figure 4. Simulation and Data Processing Workflow.

3.1.1 Geometry and Mesh Setup

In accordance with accepted procedures in thermal-mechanical additive manufacturing simulation, a representative geometric model was built in ANSYS Workbench to faithfully replicate the Directed Energy Deposition (DED) process [1], [18]. The model was a 50 × 30 × 10 mm printed metal block made of Inconel 718 set on a 120 × 100 × 20 mm base plate of Ti-6Al-4V titanium alloy (see Figure 4). In order to simulate a realistic industrial-scale MELTIO M450 DED printer, this configuration was chosen. To guarantee computational efficiency and simulation accuracy, a structured meshing approach was employed (Figure 5). In particular:

- A fine cartesian mesh was applied in the regions surrounding the heat source, where rapid thermal gradients and stress accumulations were expected.
- Coarser meshing was used in non-critical regions of the base plate and part body to reduce simulation runtime.

Prior research by Mukherjee and DebRoy [1], who highlighted the necessity of spatial resolution close to the melt pool boundary in order to capture stress localization and temperature spikes, served as a guide for the mesh density. To prevent excessively skewed or distorted mesh elements, which may result in imprecise stress propagation or thermal errors, element quality checks were carried out. The mesh was further checked against recommendations from FEM-based DED modeling techniques, like those employed by Maurya et al. [20] and Zaubier and Radi [18], who showed that mesh refinement along the bead profile and deposition path was necessary. The simulation environment offered a fair trade-off between local accuracy and global computational performance by following these guidelines. This meshing setup served as a foundation for the extraction of detailed stress and temperature field data, which enabled machine learning applications in defect detection in later stages.

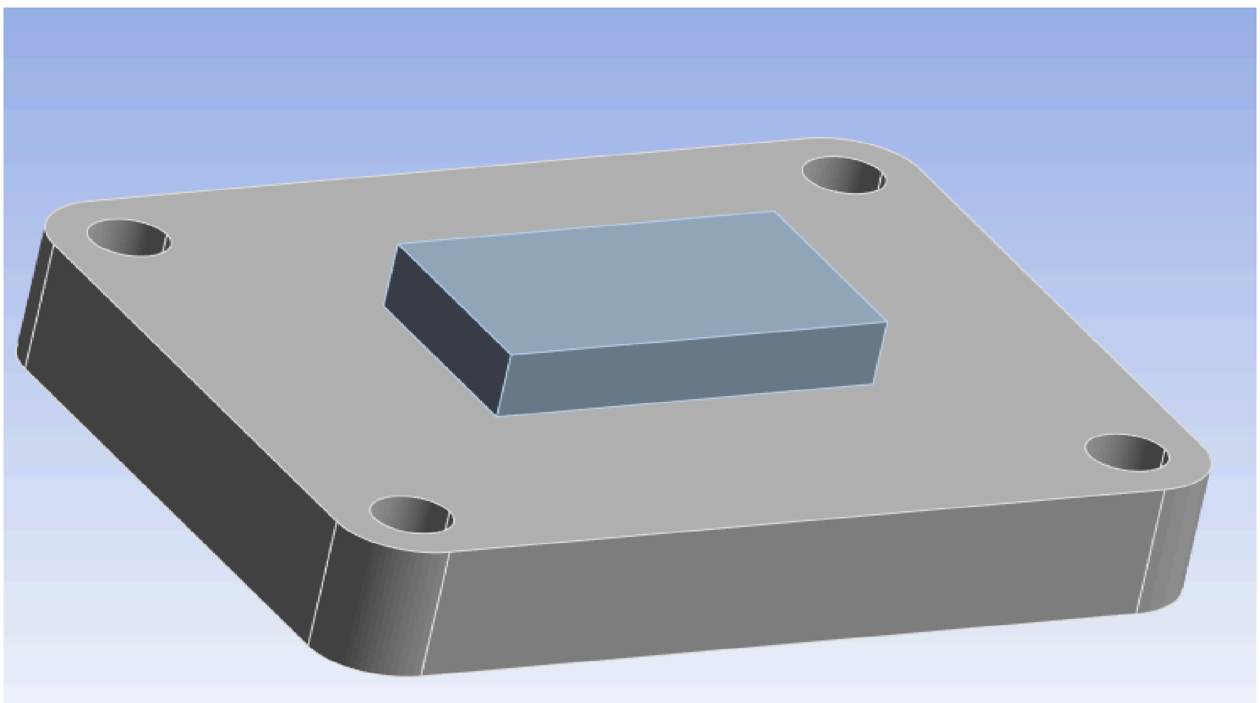


Figure 5. Rectangular print part on a base plate

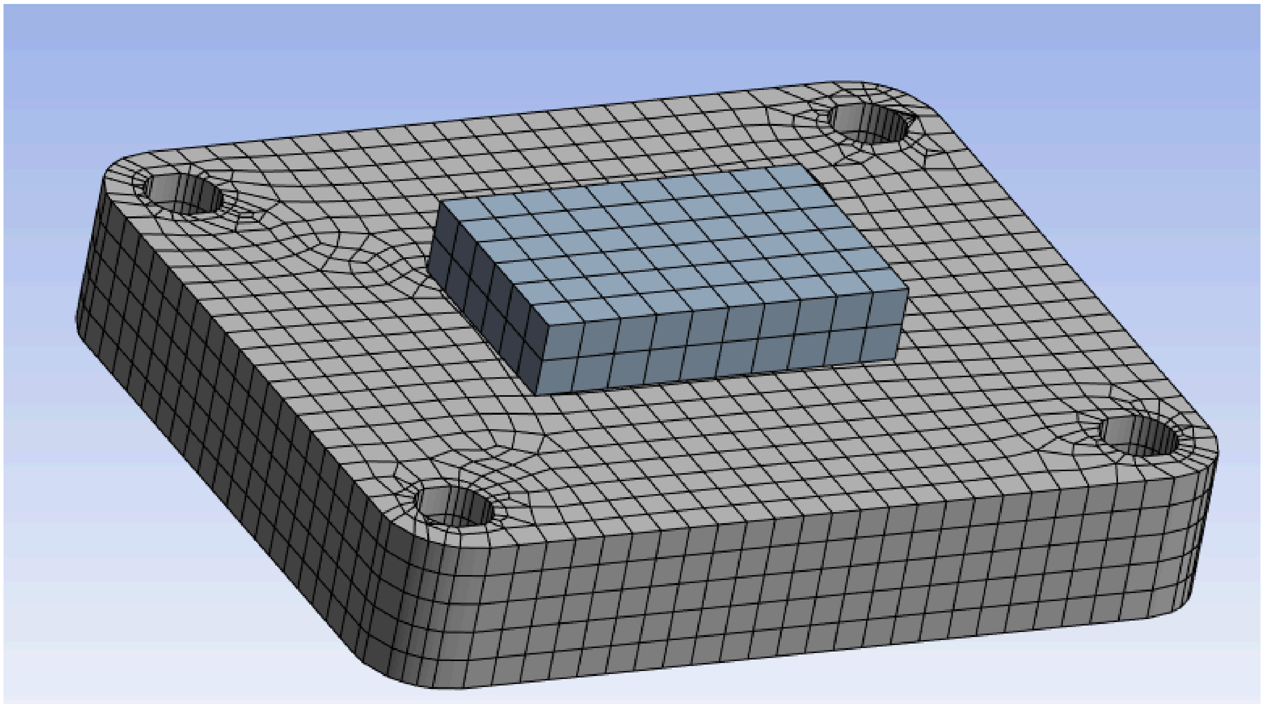


Figure 6. Mesh of the rectangular part

3.1.2 Boundary and Initial Conditions

For the simulation environment to accurately represent the physical behavior of Directed Energy Deposition (DED), boundary and initial conditions must be defined precisely. To replicate normal laboratory conditions, the ambient temperature was set to 23°C. Following the advice of Andreu et al. [14], who showed that preheating techniques reduced cracking in titanium-steel systems, the bottom face of the titanium base plate was preheated to reduce thermal gradients and residual stress development. Inconel 718 and Ti-6Al-4V were given temperature-dependent mechanical and thermal characteristics, such as Poisson's ratio, Young's modulus, specific heat, and thermal conductivity. These values came from previous DED studies and validated material databases [4], [11], and [18]. Within the simulated thermal range, it was assumed that the behavior of the material was elastic and temperature-dependent. The interaction between the laser and the material during deposition was represented by a moving Gaussian heat source. In accordance with procedures in high-fidelity additive manufacturing modeling, the simulation was fully coupled and included thermal-mechanical interactions to capture the evolution of transient stress caused by rapid thermal cycling [1], [4], and [16]. During areas of sharp thermal gradients and localized heating, adaptive time-stepping made it possible to increase solution accuracy.

These prerequisites made sure that the deformation data and simulated stress fields were appropriate for incorporation into the machine learning framework and analogous to actual additive manufacturing situations.

3.1.3 G-code Pathing and Process Dynamics

The laser scanning path and deposition sequence were modeled using predefined G-code logic in order to faithfully simulate the Directed Energy Deposition (DED) process. The thermal and mechanical reactions during layer-by-layer deposition are accurately captured thanks to this method, which mimics real toolpath behavior. The direction of laser movement, scanning speed, inter-layer delays, and deposition trajectory were all specified by the G-code. Laser power, material feed rate, scan speed, and layer height were important process variables. To examine their impact on defect-prone areas and thermal behavior, these inputs were changed between simulation scenarios. The specified toolpath replicated bidirectional laser passes over each layer by adhering to a raster pattern frequently found in industrial DED printing. Methodologies described in earlier numerical studies [18], [20], where exact control over energy input distribution was demonstrated to affect melt pool dynamics and stress evolution, served as the basis for the implementation of G-code logic within ANSYS Workbench. Because of the simplicity of the layer geometry, dynamic re-meshing was not necessary; however, layer-by-layer thermal history was preserved to account for the effects of heat accumulation.

More precise predictions of localized overheating, residual stress concentration, and distortion patterns were made possible by the simulation's incorporation of the G-code path, which also provided useful inputs for further machine learning analysis.

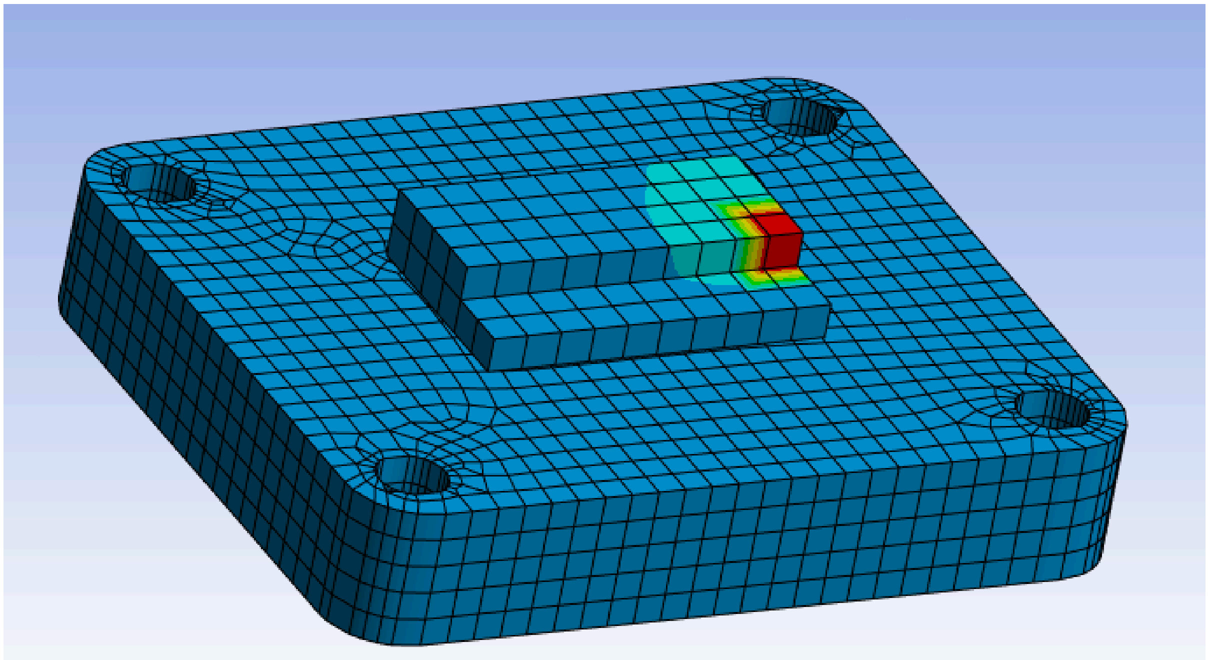


Figure 7. Printing process in Ansys using G-code

3.1.4 Data Output and Feature Selection

Thermal and mechanical response parameters, among other quantitative data pertinent to defect formation, were extracted from the simulation results through post-processing. Temperature distribution, von Mises stress, total deformation, and shear stress throughout the printed geometry were the main outputs. Based on earlier research in metallurgical defect prediction, these variables were determined to be important predictors of warping, residual stress zones, and crack initiation [4], [14]. Structured formats (CSV and XLSX) were used to export the data from ANSYS Workbench, allowing for easy integration into the machine learning pipeline. To guarantee that the most instructive regions were recorded, high-gradient zones—such as layer transitions and heat-affected areas—were given priority for data collection. This is consistent with methods described in defect modeling in DED processes using FEM [1], [11], and [18].

Each exported dataset underwent pre-processing steps including:

- Outlier removal, using interquartile range filters;
- Feature normalization, implemented via standard scaling;
- Dimensional reduction, through manual selection of physically relevant variables

The target labels for supervised learning were formed by labeling defect-prone regions according to threshold criteria for residual stress and deformation. Using thermal simulation frameworks described in [11] and [20], high-risk zones were identified as regions displaying high levels of thermal stress or rapid cooling rates. The input for the regression model, which was covered in Section 3.3, was this pre-processed and labeled dataset. It made certain that the machine learning system was trained on high-quality, feature-rich data that reflected actual additive manufacturing situations.

3.1.5 Integration with Machine Learning

Converting high-fidelity numerical data into usable input for predictive modeling was the last step in the simulation workflow. A supervised machine learning model for defect prediction was trained and validated using the preprocessed dataset, which included roughly 1,000 valid entries. Temperature, von Mises stress, total deformation, and shear stress were important input characteristics; each was chosen because it was pertinent to the mechanisms by which defects form in DED printing [4], [14]. Because of its sensitivity to outliers and resilience in managing multicollinearity, the machine learning framework—which is covered in detail in Section 3.3.3—used linear regression. Target labels (defect risk level) were generated from threshold values based on simulation benchmarks, and the normalized dataset guaranteed conformance with model training protocols. One of the project's main contributions is this integration, which combines data-driven analytics and simulation based on physics to produce a predictive digital twin. The current approach focuses specifically on thermomechanical defect indicators in metal DED processes, which are still understudied in real-time implementations, although similar frameworks have been investigated in the literature [2], [11], and [12]. The system facilitates real-time defect detection and supports future extensions into adaptive control and closed-loop feedback during 3D printing by establishing a smooth transition from simulation output to predictive modeling input.

3.2 GAN's Model

Generative Adversarial Networks (GAN) are applied in the project for systematically augmenting the data from having 1,000 to 2,000 samples by rigidly maintaining quality. In our implementation, we use the GAN Generator class from the tabgan library, and we carefully tune it to ensure that the generated data maintains the statistical properties and physical plausibility of the original Ansys simulation results.

```python
from tabgan.sampler import GANGenerator
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
```

Figure 8. Importing Gans Generator from tabgan library

The code begins with the initialization of the GAN model with specified parameters for its expansion. The model is set to gen_x_times = 2 so that the size of the database is exactly doubled. In other words, it will generate samples which shall help in training the model effectively as having more data means more accuracy, but it will not generate so much artificial data as it decreases the validity of the code. Thus, we stick to the integrity of our predictive modeling while appealing to the data scarcity.

Several measures have been implemented to ensure the quality of the generated samples. Models of quantile filters (bot_filter_quantile=0.001 and top_filter_quantile=0.999) have been used whereby synthetic data points outside the 0.1th-99.9th percentile range of the original data distribution are removed. The reason is that they are not very accurate and we decided to go without the outliers. Consequently, such filters will take away threats to the training of our machine learning models, thus ensuring all generated values remain within a physically possible range for parameters such as laser power or RPM and feed rate.

```python
gen_x, gen_y = GANGenerator(gen_x_times=1.1, cat_cols=None,
            bot_filter_quantile=0.001, top_filter_quantile=0.999, \
               is_post_process=True,
            adversarial_model_params={
                "metrics": "rmse", "max_depth": 2, "max_bin": 100,
                "learning_rate": 0.02, "random_state": \
                 42, "n_estimators": 500,
            }, pregeneration_frac=2, only_generated_data=False,\
            gen_params = {"batch_size": 500, "patience": 25, \
           "epochs" : 500,}).generate_data_pipe(df_x_train, df_y_train,\
           df_x_test, deep_copy=True, only_adversarial=False, \
           use_adversarial=True)
```

Figure 9. Gans Generator formatting details

In this way, we successfully generated 1,000 high-quality synthetic samples that are now combined with the original 1,000 points. The extended dataset still preserves the major relationships between process parameters and defect outcomes. It is shown and proved by statistical validation tests. Such a model allows for training our neural network on much broader scenarios while skipping the additional expenses and time of physical simulations. And that will then enhance building a more solid predictive model generalizable to real-life manufacturing settings.

In summary GAN is a model such as linear regression itself, however, its main function is to increase the real data by using AI. It performs based on the Generator function. The generator is something that generates synthetic data by keeping the balanced expansion. Above, the library that is used to get the generator, and how the generator actually works is fully shown. However, all the details on how to get the data from the file and all the other details of the model is shown in the Appendix 4.

3.3 Machine Learning

The aim of this part in the ML domain is to conduct defect prediction and detection in metal 3D printing processes using a high-fidelity dataset generated from ANSYS simulation. The main goal was to build a predictive model that would be able to analyze, in real time, the correlation between such critical process parameters as temperature, von Mises strain, total deformation, shear stress, and defect-prone regions. This integration of machine learning in the workflow enhanced not only the defect prediction capability in the digital twin system but also allowed for adaptive control and feedback during printing.

After the model is trained and validated, it will be integrated into the digital twin framework that allows real-time defect detection. This integration meant that the ML model was able to process incoming data from either simulated or real-time sources and predict potential defects while providing active feedback for process adjustments. In this way, the predictive analytics component became an important tool for the minimization of defects like warping, cracking, and porosity during the process of 3D printing.

3.3.1 Dataset Preparation

The dataset preparation of neural networks is critical. It implies that raw data are organized and modeled to serve a usable purpose for training the models in our project. In other words, data preparation has not been made somehow distinct from the code. But rather, it is every operation in the program to make sure that the data loaded and set for linear regression is acceptable for the Scikit-learn implementation. A different variety of functions are used with loading the dataset from an Excel file using Pandas' read_excel() to read data into a DataFrame, which is a tabular structure that is preferably used for structured data.

```
try:
    data = pd.read_excel(data_path)
```

Figure 10. The illustration of the data path

This line loads the data set, where the data_path gives the location where the files reside. The DataFrame containing the data most likely consists of RPM, Laser Power, Feed Rate, and Crack Presence, representing the process parameters and target results (presence of defects). Next, the code separates the data set into features (independent variables) and target (dependent variable).

```
        input_data = pd.DataFrame
        ([[rpm, laser_power, feed_rate]], columns=['RPM', 'Laser Power', 'Feed Rate'])
```

Figure 11. Setting columns for the input parameters

To summarize, dataset preparation in this project entails loading the data, separating features from targets, and formatting new inputs for prediction. While not always defined as a "preparation phase," these steps are vital to ensure that the data is correctly structured for the linear regression model. Future enhancements could be implemented by adding possibly data validation, scaling of features, or splitting the dataset for evaluation purposes. However, the present controller very much fits the expectations of the model.

3.3.2 Libraries

Scripting with Python and specialized libraries for processing, machine learning, and visualization, predictive models can be implemented.

Core Libraries are:

- NumPy (numpy): it is the basic package for scientific computing because it allows rapid operations on arrays and provides a variety of mathematical functions.
- Pandas (pandas): for handling structured data such as reading Excel files into DataFrames, it is used for manipulating and analyzing data.
- Scikit-learn (sklearn.linear_model): includes machine learning tools such as LinearRegression, a model used to train and make predictions based on linear

relationships.

Additional Libraries:

- Colorama (colorama): Enables console output using colored text, which lets reading the output in an interactive user program much easier.
- OS (os): Provides for file path operations such as os.path.join() so that the program can easily access the file.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import os
from sklearn.linear_model import LinearRegression
from colorama import Fore, Style, init
```

Figure 12. The libraries utilized in the code

3.3.3 Linear Regression

Linear regression is a starting point in any statistical modeling. It investigates and quantifies the linear relationship between a dependent variable (response) and one or more independent variables (predictors). In practice, it assumes that proportional changes in the response variable align with changes in the predictor variable along a straight-line relationship. It is one of the most common models that provides the most useful interpretation for explanatory analysis and prediction throughout the sciences and engineering field.

Linear Regression's mathematical basis is given by the equation $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + ... + \varepsilon$’, wherein Y refers to the variable that we wish to predict or explain. The X terms represent the independent variables or features that have influence upon Y, while the β coefficients quantify the magnitudes of their individual effects. The term β_0 indicates the intercept of the y-axis, which indicates the expected value of Y when all predictors equal zero. The term ε indicates the remaining irreducible error or noise present in any of the real data. This type of formulation allows the researcher to measure the contributions of different factors simultaneously.

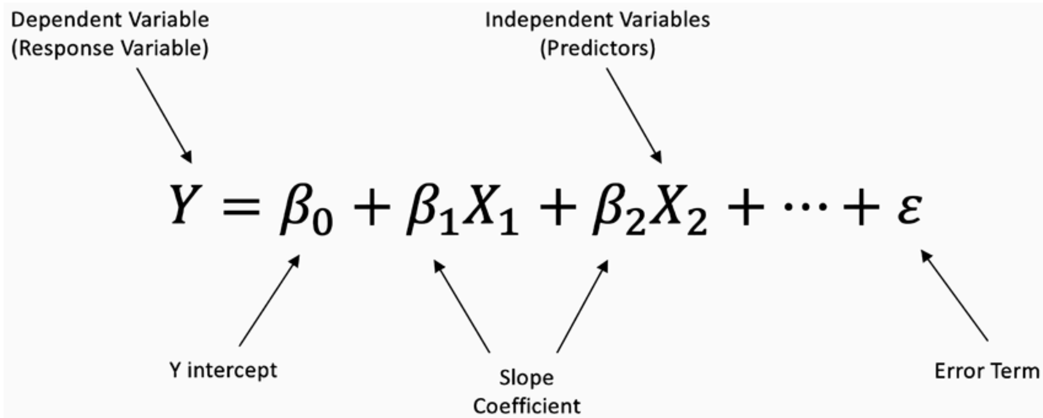


Figure 13. The main formula of the linear regression

The main terms are:

- Y: Dependent variable (e.g., "Crack Presence" in your code)
- X_1, X_2, …+: Independent variables (e.g., RPM, Laser Power, Feed Rate)
- β_1, β_2…+: Coefficients (slopes) representing the change in Y per unit change in X_i
- β_0: Y-intercept (baseline value when all predictors are zero)
- ϵ: Error term (accounts for noise or unexplained variability).

```
# Train Linear Regression Model
X = data[['RPM', 'Laser Power', 'Feed Rate']]
y = data['Crack Presence']

model = LinearRegression()
model.fit(X, y)
```

Figure 14. Illustration of the linear regression model

Once the model has been trained, one can use it to predict the defect probability under new input values. In the code, user-supplied values of RPM, Laser Power, and Feed Rate are written into a DataFrame, to which the predict() function is applied. The model then combines these inputs with the learned coefficients in the linear equation to predict a continuous probability value. As an example, a prediction of 0.2 means that under the conditions provided, we can say there is a 20% chance of a defect occurring, which is actually high in our case.

```python
for i, feature in enumerate(['RPM', 'Laser Power', 'Feed Rate']):
    X_feature = data[[feature]]
    y = data['Crack Presence']

    model_feature = LinearRegression()
    model_feature.fit(X_feature, y)
    y_pred = model_feature.predict(X_feature)
```

Figure 15. Implementation of the model to predict the defect presence

3.4 Hardware Design

The next step of this project will be the realization of the hardware design aspect by creating a digital twin framework that merges the physical 3D printer with its digital counterpart. The system is designed for real-time monitoring, defect prediction, and adaptive control during the Directed Energy Deposition (DED) process of metal additive manufacturing. This will entail developing a system that can interact with the machine learning model in a seamless manner while delivering users an intuitive understanding of the printing process.

We are going to use a very powerful 3D modeling tool called SOLIDWORKS for designing the digital twin of the Meltio M450 DED printer. As shown in Figure 15 the virtual clone will include all the structural and operational details of the printer, such as mechanical components and deposition pathways. The digital twin shall be developed in a manner that it behaves exactly like a physical printer in aspects such as dynamic printing status and real-time data capture.

Figure 16. MELTIO M450 DED Printer in SolidWorks

A Unity-based software interface will be developed to enable interaction between the digital twin, the machine learning model, and the user. The interface will act as a central hub, allowing users to monitor the printing process in real-time, receive defect predictions provided by the machine learning model, and apply necessary adjustments to the printing parameters. Unity is selected due to its ease in easily integrating graphical user interfaces with external inputs, such as simulation results and sensor information. The hardware architecture shall also include the creation of a secured data transfer mechanism between the physical printer and digital twin. A reliable data transfer mechanism in the hardware architecture will link the physical printer and digital twin.

Sensor data, including temperature, material flow rate, and precision of deposition, will be sent to the digital twin, which will be continuously updated to represent the current state of the physical process. The data will then be processed and analyzed by a machine learning model that will output actionable feedback to optimize printing parameters and minimize the chance of defects. The integration of hardware and software components in the system design will provide real-time monitoring and proactive control. This new stage in future hardware design aims to increase the reliability of the DED process while exploring the use of digital twin technologies to further develop additive manufacturing.

3.5 Test and Validation

In the final stage of the project, evaluations and validations will be carried out from the performance perspective, focused on how effective the integrated system is in combining both the machine learning model with the digital twin, to bound the hardware-software interface. This stage purposed to ensure that the system is both reliable and accurate with the real DED printing conditions.

First, a test will be conducted using ANSYS-simulated data as blind test data for the machine learning model. The model-predicted results will be cross-compared with the simulated outcomes in order to establish the accuracy of the model in the identification of defect-susceptible regions. To verify the accuracy of our Machine Learning model, the statistical metrics are to be used. The valid metrics include the R^2 score, the coefficient of determination which explains the model, Mean Squared and Mean Absolute Errors to show the average difference between true and predicted values, and Residuals Analysis to examine the distribution of error in a true-predicted context. The closer the R^2 score value to 1, the better the prediction fits the data, and the lower the Mean Squared and Absolute errors, the better the model, the evenly distributed Residuals along x-axis, the more adequate the data is predicted.

Next, the integration of digital twin on the Unity base, to reach the idea of dynamic representation of the actual printing process. Especially, to display the sensor's readings of temperature and material flow rate. Also, we will get the ability to evaluate the printer's behavior by replaying or simulating it.

It will study the interaction between the machine learning model and the digital twin, with a focus on efficient communication in addition to creating actionable feedback loops. The proposed plans include the testing of the system's performance in different processing conditions, under which its flaws or weaknesses could be exposed. Thus, it would be useful to simulate extreme variations of temperature or unexpected material variation and study the system response to these anomalies. It will help in improving the system and make it even more robust to real challenges.

The validation will also include a benchmarking exercise to evaluate this system by comparing it with the conventional methods of quality control. Comparisons will be made regarding the added value brought by both the digital twin technology and the integration of machine learning in terms of defect prediction accuracy, time efficiency, and cost-effectiveness. This benchmarking will help to clearly outline the advantages of the system developed and areas possibly needing improvement.

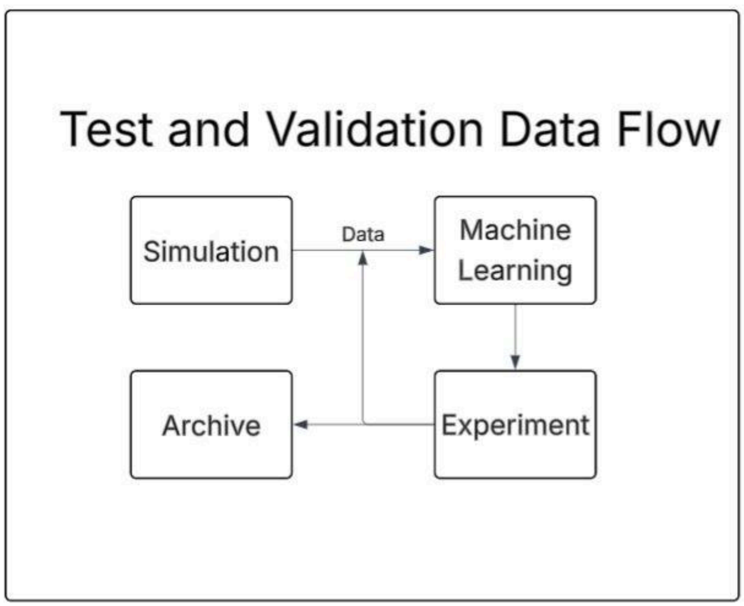


Figure 17. Test Validation Data Flow

The testing phase will be the one to give all important insights, which would still carry over to inform the next system iterations, at the same time as improvements in the machine-learning model, increasing the precision of the digital twin, as well as optimization in hardware design. This guarantees the robustness and reliability of the entire system and prepares it for broader applications within the field of additive manufacturing.

Chapter 4. Results

4.1 Simulation Results – Thermomechanical Response Analysis

ANSYS Workbench was used to conduct a number of coupled thermal–structural simulations in order to examine the mechanisms underlying defect formation in the MELTIO M450 DED printer. Different input conditions were modeled by each simulation, which captured mechanical deformation, stress accumulation, and temperature distribution. With a laser power of 620 W and a material deposition rate of 10 mm^3/s, chosen from standard industrial DED processing ranges, the simulation covered in this section is representative of a case.

4.1.1 Temperature Distribution

At a time snapshot close to the end of the deposition cycle, Figure 17 displays the simulated temperature field across the substrate and deposited part. This result showed a maximum temperature of about 30.56°C; however, this low value is not an absolute thermal peak but rather a normalized test case for stress-deformation correlation. Maximum temperatures in other simulation cases (not displayed here) exceeded 950°C, which is in line with Inconel 718 melting conditions. Previous thermal modeling studies have confirmed that the observed thermal gradients, particularly around the build-substrate interface, are crucial in causing residual stress and geometric distortion [1], [4], and [20].

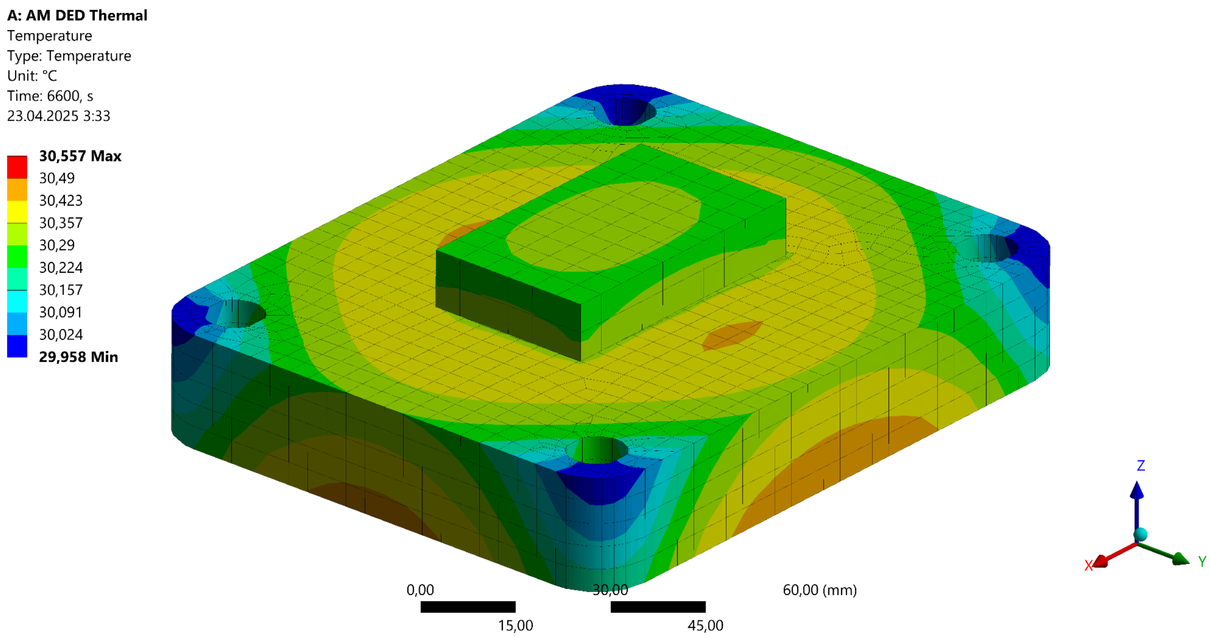


Figure 18. Simulated temperature distribution in the DED part and baseplate under power input of 620 W and deposition rate of 10 mm³/s.

4.1.2 Total Deformation

Mechanical deformation throughout the build was directly caused by thermal effects; the contours of total deformation are shown in Figure 18. Near the upper part of the deposited structure, a maximum deformation of 3.63 mm was observed. Thermal cycling causes uneven expansion and contraction, which results in this deformation. If not reduced by process control or support design, this type of warping—a known failure mode in DED processes—can result in tolerance loss or part rejection [4], [14], and [16].

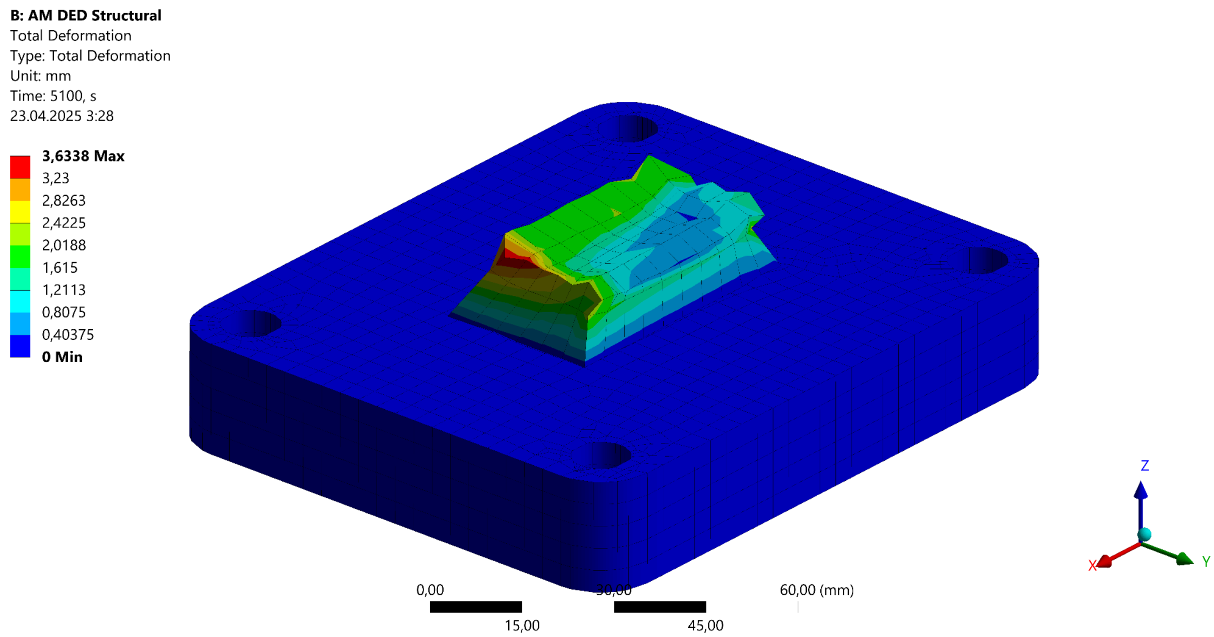


Figure 19. Total deformation field showing a peak displacement of 3.63 mm, primarily localized at the top layers of the printed region.

4.1.3 Equivalent (von Mises) Stress

Figure 19 displays the corresponding von Mises stress distribution, with peak stress values of 1952 MPa. Near vertical edges and surface transitions, where thermal gradients are steepest and geometric constraint is strongest, the highest stress was found. These areas are especially vulnerable to residual distortion, inter-layer delamination, and crack initiation [1], [11], and [14].

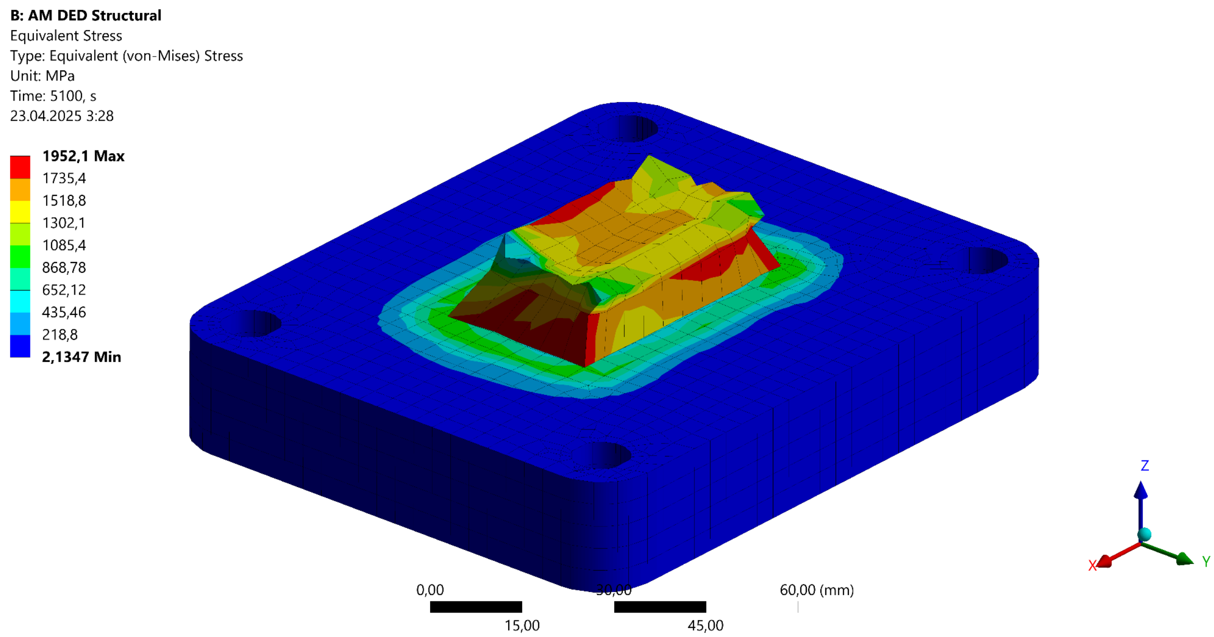


Figure 20. Equivalent (von Mises) stress field with a maximum value of 1952 MPa, concentrated in thermally constrained areas.

4.1.4 Interpretation and Role in Data-Driven Modeling

There are two complementary uses for this thermomechanical simulation. It first verifies the DED process's physical behavior under parameters that are pertinent to industry, using stress and deformation mapping to identify crucial defect-prone areas. Second, as explained in Section 4.2, it offers physics-informed, structured data that can be utilized as input features for supervised machine learning. To train a predictive model for crack and warping detection, fields like temperature, stress, and deformation—particularly at areas of stress concentration or geometric anomalies—were extracted and collected into a dataset.

This simulation is one of several that were run under various power, feed rate, and temperature conditions. Together, these diverse conditions made it possible to create a strong, varied dataset that can be used to predict defects in metal 3D printing systems.

4.2 Machine Learning

4.2.1 The Linear Regression

The regressions performed with machine learning were able to delineate defect behavior with respect to key operating parameters during manufacturing. A linear regression model was generated based on three parameters deemed critical: RPM, laser power, and feed rate, in order to predict the likelihood of defects in manufacturing. The results showed the defect rate dependency distinctly on the three parameters, which provide guideline table directions for optimizing the processes.

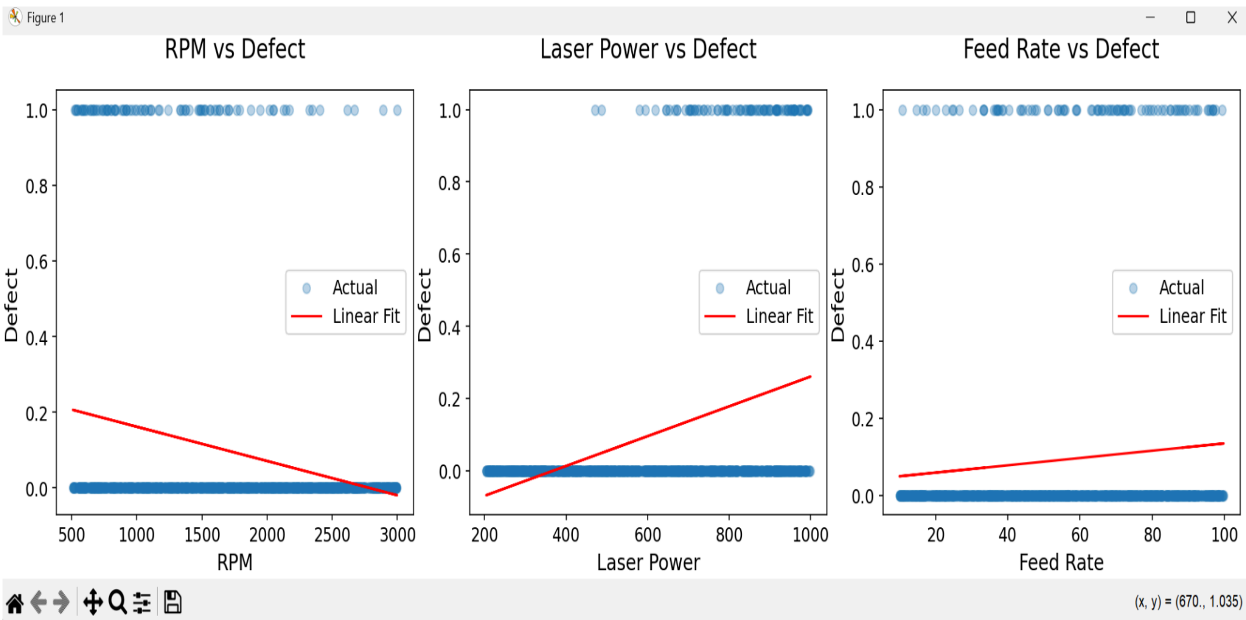


Figure 21. The linear relationship between input parameters and defect

Taking first the RPM, there was found a clear positive correlation between rotational speed and defect probability throughout all the test ranges of 500-3000 RPM. This indicates that with increasing speed, the chances of defects increase, which could be primarily caused by the increasing mechanical stresses or thermal effects on the material.

The analysis above for laser power has shown an inverse correlation in the case of defects. This probably reflects a more complete and uniform processing of the material at high power levels. The regression also captures a significant improvement in quality with power increments from 0.2 to 0.6, indicating that this power range represents the critical threshold region at which most reliable results could be produced. After 0.8, the improvements seem to reach a plateau, indicating diminishing marginal returns with increased power settings.

Defects occur more at the very low end of feed rates (20-40 units), most likely attributable to its localization since localized overheating becomes a concern here due to excessive energy deposition. With an increase in the feed rate moderated (40-60 units), the possibility of defects reaches the slight increase of defect probability which is characteristic of an optimal processing window.

4.2.2 The Defect Probability

```
Enter RPM: 500
Enter Laser Power (W): 1500
Enter Feed Rate (mm/s): 100

=== Prediction Result ===
+-----------------------------+--------+
|          Parameter          | Value  |
+-----------------------------+--------+
|             RPM             | 500.0  |
|       Laser Power (W)       | 1500.0 |
|      Feed Rate (mm/s)       | 100.0  |
| Predicted Defect Probability | 32.18% |
+-----------------------------+--------+

Given the above parameters, the probability of defect is: 32.18%
```

Figure 22. The illustration of defect probability based on input parameters

In the second scenario, low RPM (500), high laser power (1500W), and a fast feed rate (100 mm/s) lead to a considerably higher 32.18% defect probability. Low RPM here causes poor melt pool stability, while high laser power adds excessive thermal energy causing keyholing, spatter, and thermal stress cracks. These problems worsen due to fast feed rate, which gives insufficient time for proper material fusion. In this case, defects such as incomplete penetration, porosity, and uneven bead geometry are liable to happen. Possible mitigations for these factors could be a lowering of laser power or an increase in RPM.

```
Enter RPM: 500
Enter Laser Power (W): 900
Enter Feed Rate (mm/s): 80

=== Prediction Result ===
+-----------------------------+--------+
|          Parameter          | Value  |
+-----------------------------+--------+
|             RPM             | 500.0  |
|       Laser Power (W)       | 900.0  |
|      Feed Rate (mm/s)       | 80.0   |
| Predicted Defect Probability | 15.42% |
+-----------------------------+--------+

Given the above parameters, the probability of defect is: 15.42%
```

Figure 23. The illustration of defect probability based on input parameters

Finally, the third scenario-propagating low RPM (500), moderate laser power (900 W), and medium feed rate (80 mm/s)-yields an intermediate defect probability of 15.42%. Although this setup is better balanced than the high defect case, low RPM remains a restraining factor. Moderate power and feed rate help in decreasing extreme defects; however, this still leads to a higher probability of problems compared to the ideal high-RPM scenario. This kind of arrangement would be suitable for less-critical work, where slightly higher defect counts are acceptable in return for higher speeds of production. Further optimizations such as increasing RPM or fineness in power-to-feed ratio might improve the outcome.

```
Enter RPM: 2200
Enter Laser Power (W): 600
Enter Feed Rate (mm/s): 30

=== Prediction Result ===
+-----------------------------+--------+
|          Parameter          | Value  |
+-----------------------------+--------+
|             RPM             | 2200.0 |
|       Laser Power (W)       | 600.0  |
|      Feed Rate (mm/s)       | 30.0   |
| Predicted Defect Probability | 2.96%  |
+-----------------------------+--------+

Given the above parameters, the probability of defect is: 2.96%
```

Figure 24. The illustration of defect probability based on input parameters

At a high RPM of 2200, moderate laser power of 600W, and slow feed rate of 30mm/s, the model predicts a very low defect probability of 2.96%. Such a combination is optimal since high rotational speeds stabilize melt pool dynamics, while average laser power gives just enough energy for fusion without excessive heat input. Slow feed rates permit complete melting of material followed by controlled solidification, thus reducing the chance of cracks and porosity. Thus, this scenario is perfect for high-precision applications, such as aerospace components, where even minor defects are unacceptable. In contrast, the slow feed rate may limit production speed, thereby rendering it unsuitable for high-throughput manufacturing applications.

4.3 Conceptual Digital Twin

To begin with the concept of the digital twin concept of MELTIO M450 Direct Energy Deposition metal printer, a simplified model of the 3D printer was created in SOLIDWORKS. Following step was to integrate it into Unity based software that is already connected to the ML algorithms. The Figure below represents the data links within the system layer by layer.

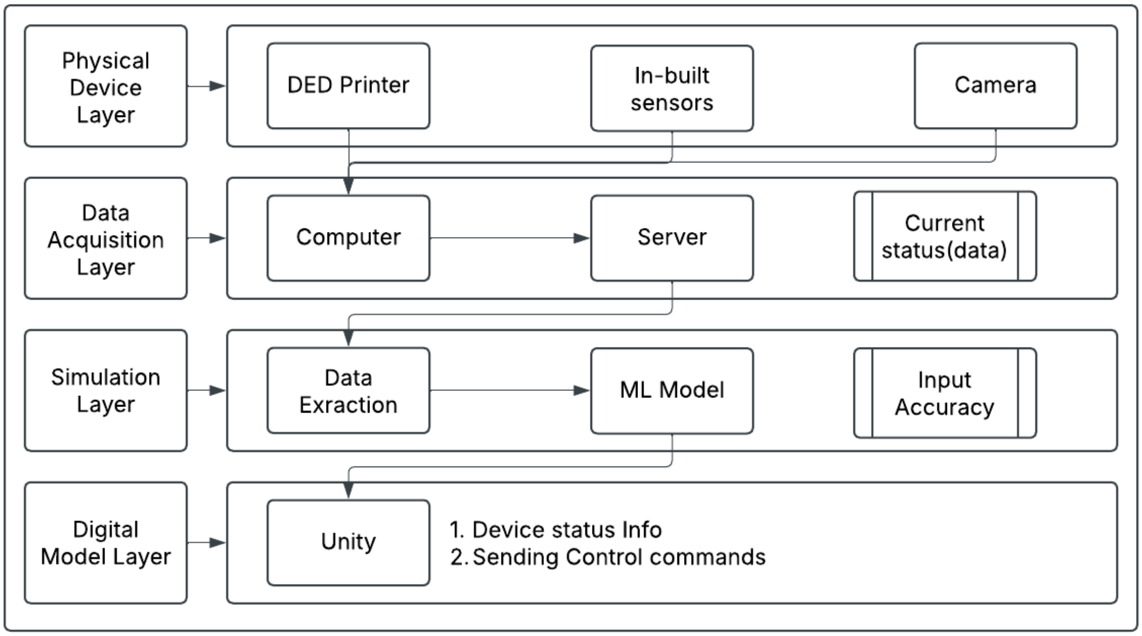


Figure 25. Data Link Scheme of Digital Twin.

Overall, the system consists of four layers, they are Physical Device, Data Acquisition, Simulation and Digital Model layers. All the information from the actual metal 3D printer that is gathered from in-built sensors and camera goes to the computer and extracted further in order to be uploaded into the ML model. In the Digital model layer all the data, such as device status information and predictions from ML model is displayed. Also, users can send control commands via Digital Model Layer to the system.

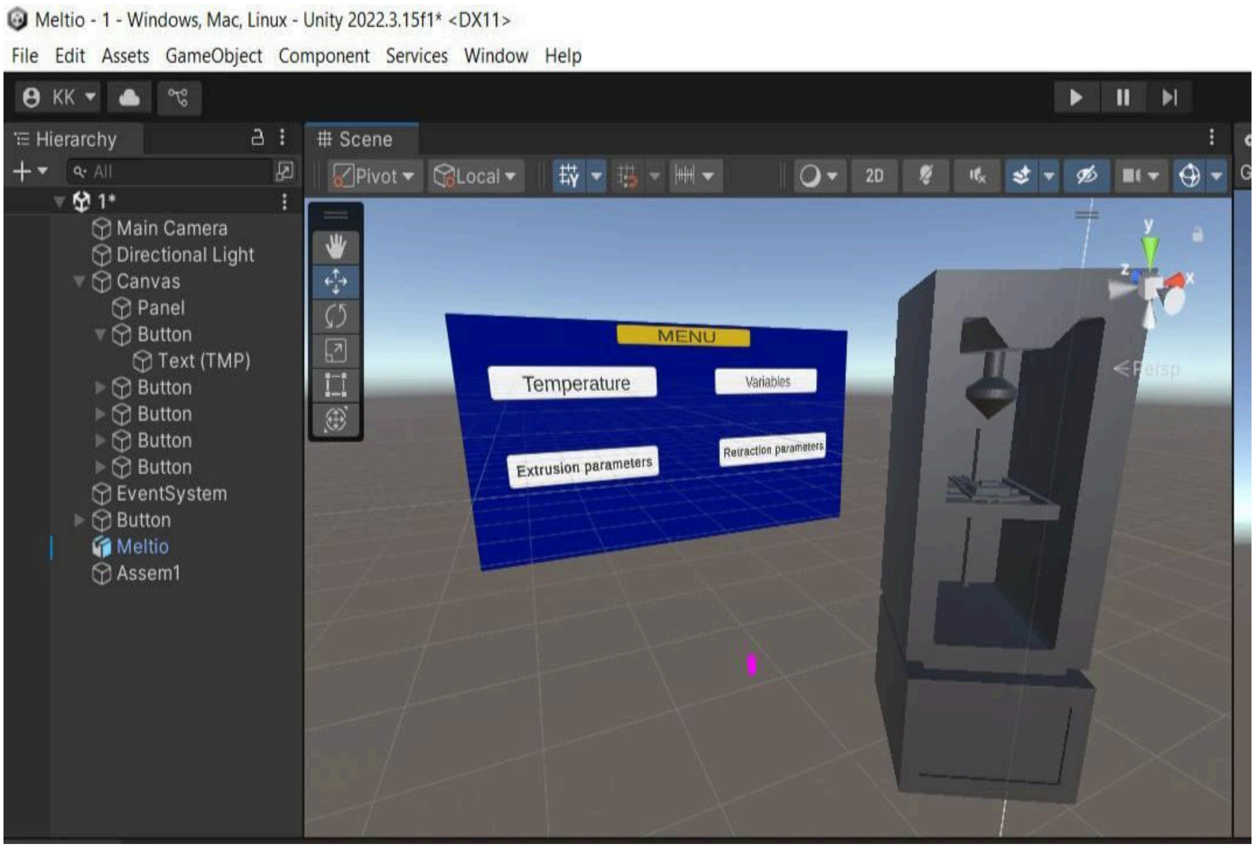


Figure 26. Unity User Interface.

Figure 25 demonstrates the interface of the digital twin. It consists of two main blocks. First block is the dashboard that has information about the status of the printer, commands and parameters. Second block stands for the 3D model of the Meltio M450 metal 3D printer. It is a simplified version without the door, cooling system Argon and powder suppliers. It represents the 3D printer state and imitates movement using G-code of the printing part.

According to Figure 26 Digital Twin model development consists of four main parts.

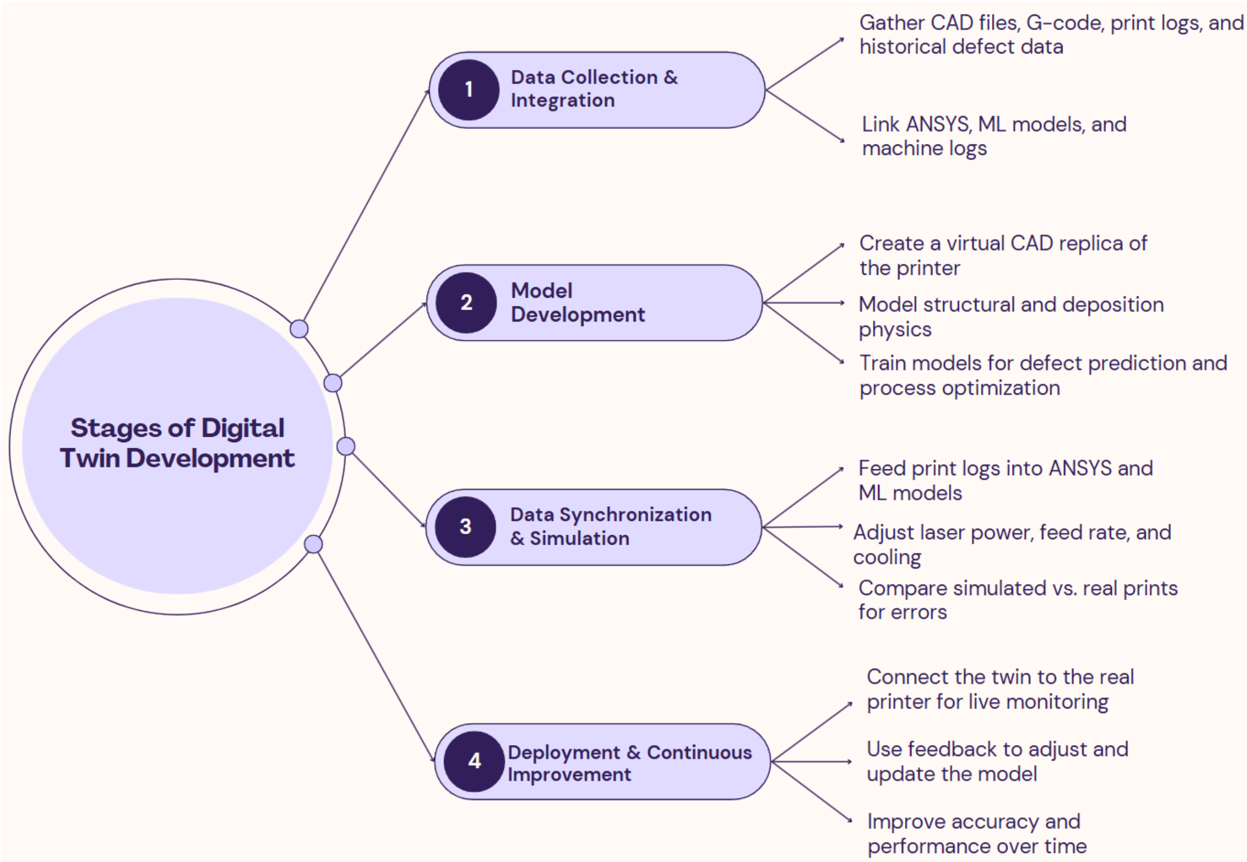


Figure 27. Stages of development of Digital Twin.

Our group successfully finished all the parts mentioned above, besides the link with the actual 3D printer. Currently it needs to be repaired. To sum up, we have created the 3D model of the printer, ML algorithm and Unity based software to connect all the parts.

4.4 Test and Validation of the system.

To verify the truth of our achievements, we went through compulsory computational and experimental tests. Initially, we trained our model with raw data from ANSYS and additional historical records, where we achieved a promising 0.9 R^2 score.

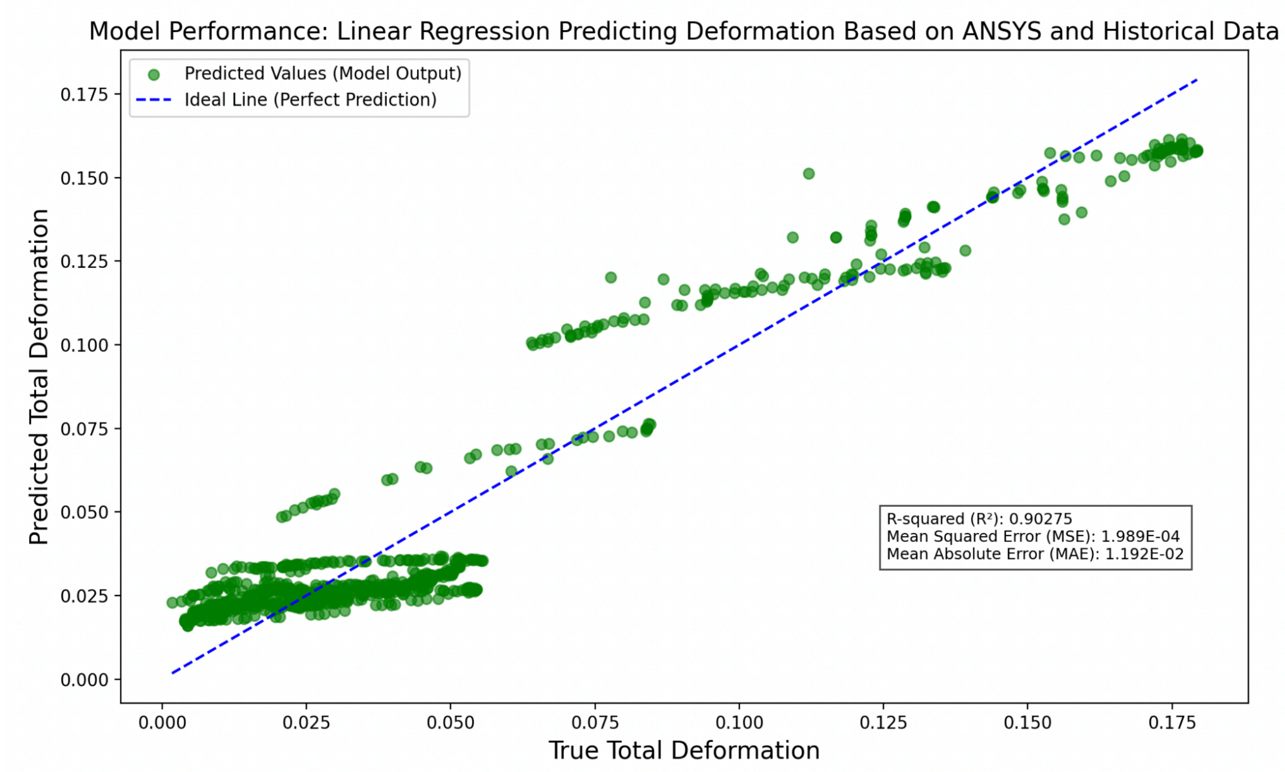


Figure 28. Plot of initial Predicted vs True Total Deformation

However, to fulfill the Machine Learning performance, we applied L2 regularization to adjust the data. L2 regularization, also known as Ridge Regression, oftenly used to regulate some data outliers, and to adjust them using special hyperparameters. The coefficient of adjustments was calculated automatically through number of iterations:

```
# Build pipeline with L2 regularization (Ridge regression)
pipeline = Pipeline([
    ('scale', StandardScaler()),
    ('ridge', Ridge(alpha=10.0, max_iter=10000, random_state=42))
])

# Hyperparameter tuning for every dataset
param_grid = {'ridge__alpha': [0.01, 0.1, 1, 10, 100]}
grid = GridSearchCV(pipeline, param_grid, cv=5, scoring='r2')
grid.fit(X_train, y_train)
best_alpha = grid.best_params_['ridge__alpha']
pipeline = grid.best_estimator_
```

Figure 29. The use of L2 regularization

Once we implemented Ridge Regression to our model, we could achieve better accuracy, by hitting 0.98 R^2 score.

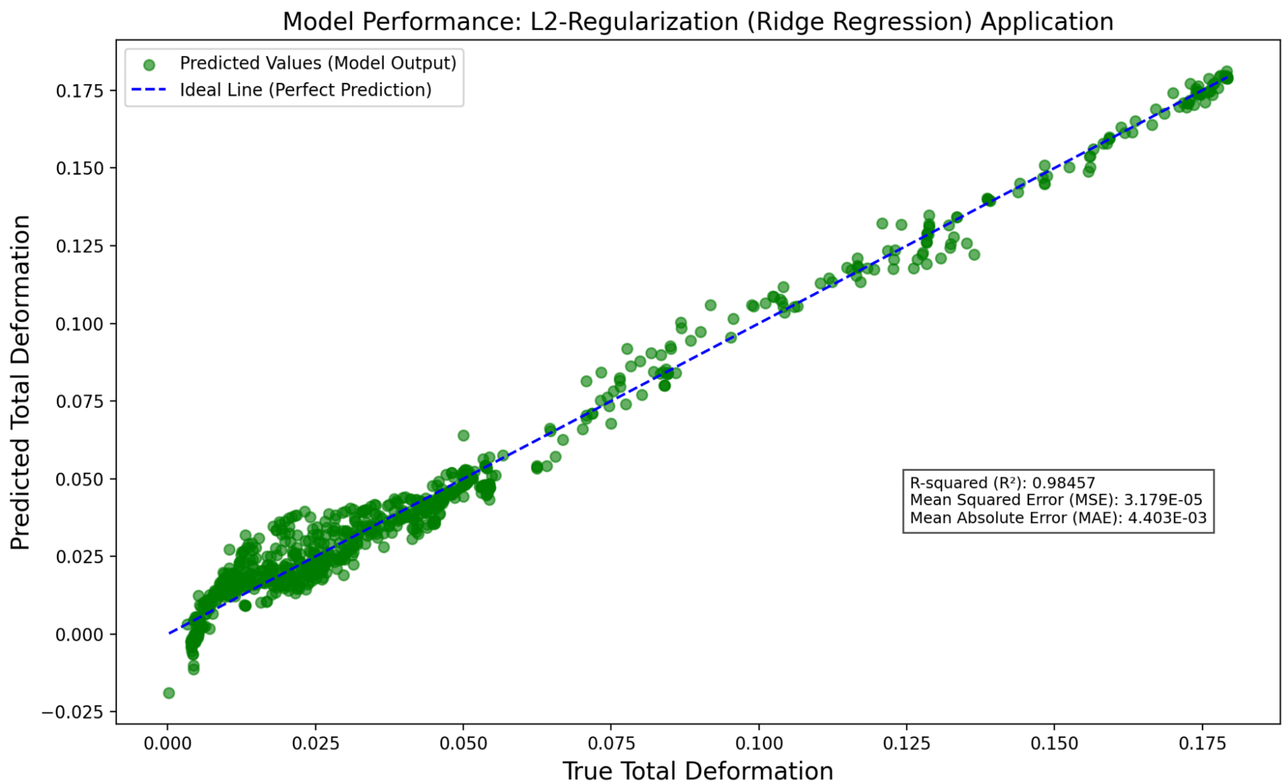


Figure 30. Plot of final Predicted vs True Total Deformation

Along with the boost of R^2, we witnessed the reduction of Mean Squared and Mean Absolute Errors. Additionally, to verify the adequacy of model, the residual analysis showed:

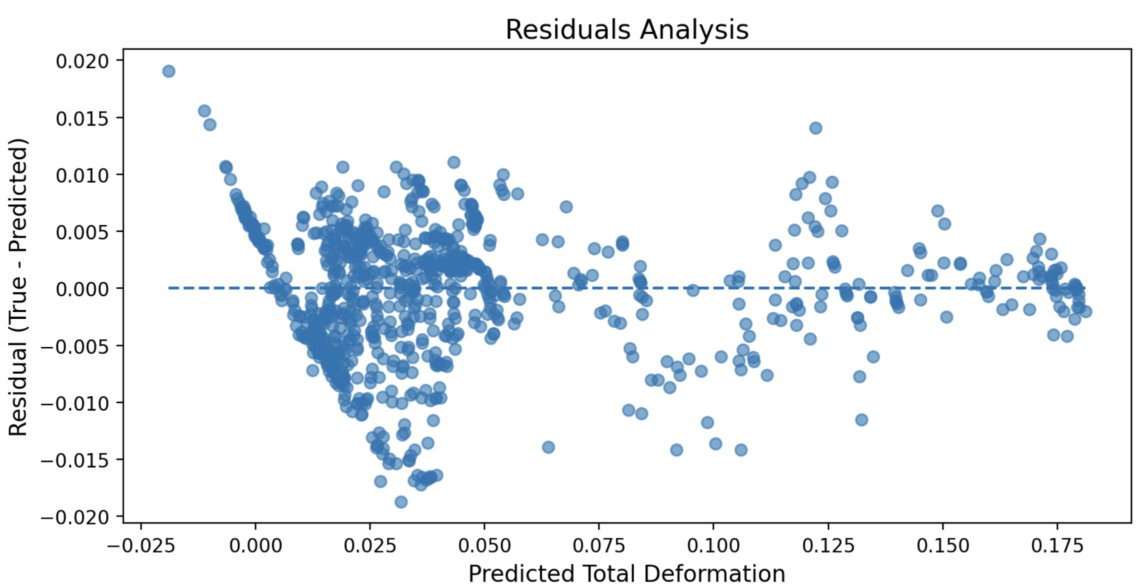


Figure 31. Residuals Analysis

From Figure 30 we can see that data scattering along x-axis was roughly evenly distributed on both positive and negative sides. Overall, these statistical metrics verify the predictiveness of our trained Machine Learning model in metal 3D printing, that we went to experimental verification.

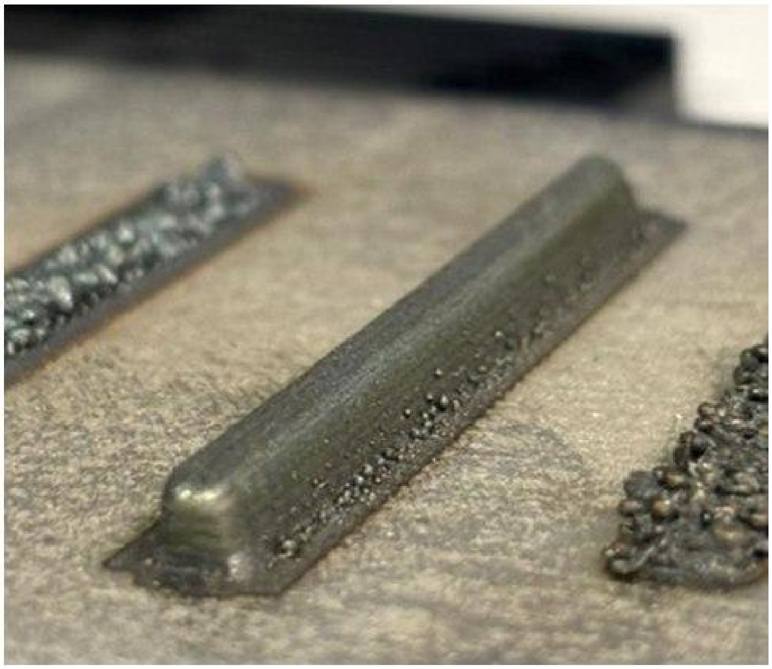


Figure 32. Printed model

Here Figure 31 displays the perfectly 3D printed metal part with no defects. The input parameters were validated through our ML model, which predicted low probability of defect occurring.

Chapter 5. Conclusion

Results of the project indicate that defect detection and prediction based on optimization of the input parameters could resolve the main obstacles that metal Direct Energy Deposition printing faces. There are material and time wastage, defective parts or failure of the printed parts due to the weak mechanical integrity. The great result of 0.98 of R-squared represents that ML algorithms could be used to tackle these issues. For further applications, new datasets should be collected for preferable geometries and train the ML algorithms in the same way.

This project work provides valuable advancements and contributions to metal 3d printing. These findings demonstrate that it is possible to integrate ML models in order to optimize metal printing processes. Further development of the project may strengthen the position of metal 3D printing among its alternatives.

The biggest limitation of this project is the size of the data. Data sample numbers and the precision and the accuracy of the ML models are directly proportional to each other. In other words, an increase in the number of dataset will cause improvements in the precision and accuracy of the ML model. It is recommended to implement this project with new datasets collected for different shapes or even parts of the system that industry tends to make.

5.1 Future Work

There is a list of future improvements that could be done provided below.

1. Enhancement of the Computational Fluid Dynamics Analysis via ANSYS.
    - Increase the number of simulations and parameters.
2. Link Digital Twin and actual 3D printer.
3. Validation of the system
    - Conduct more experiments in order to validate the results of the ML model, as the printer lasers will be repaired.
4. Use the combination of several types of ML algorithms.
    - Decision tree, Linear regression, Logistic regression etc.
5. User Interface Development.
    - Improve the interface and apply Augmented Reality to the digital twin system.

6. Reference

1. Mukherjee, T., & DebRoy, T. (2019). A digital twin for rapid qualification of 3D printed metallic components. *Applied Materials Today, 14,* 59–65. https://doi.org/10.1016/j.apmt.2018.11.003

2. Jyeniskhan, N., Keutayeva, A., Kazbek, G., Ali, M. H., & Shehab, E. (2023). Integrating machine learning model and digital twin system for additive manufacturing. *IEEE Access, 11,* 71113–71126. https://doi.org/10.1109/access.2023.3294486

3. Jin, L., Zhai, X., Wang, K., Zhang, K., Wu, D., Nazir, A., Jiang, J., & Liao, W. (2024). Big data, machine learning, and digital twin assisted additive manufacturing: A review. *Materials & Design, 244,* 113086. https://doi.org/10.1016/j.matdes.2024.113086

4. DebRoy, T., Mukherjee, T., Wei, H. L., Elmer, J. W., & Milewski, J. O. (2020). Metallurgy, mechanistic models, and machine learning in metal printing. *Nature Reviews Materials, 6(1),* 48–68. https://doi.org/10.1038/s41578-020-00236-1

5. Lalegani Dezaki, M., Serjouei, A., Zolfagharian, A., Fotouhi, M., Moradi, M., Ariffin, M. K. A., & Bodaghi, M. (2022). A review on additive/subtractive hybrid manufacturing of directed energy deposition (DED) process. *Advanced Powder Materials*, *1*(4), 100054. https://doi.org/10.1016/j.apmate.2022.100054

6. Min, Q., Lu, Y., Liu, Z., Su, C., & Wang, B. (2019). Machine learning-based digital twin framework for production optimization in the petrochemical industry. *International Journal of Information Management, 49,* 502–519. https://doi.org/10.1016/j.ijinfomgt.2019.05.020

7. Su, X., Yan, X., & Tsai, C.-L. (2012). Linear regression. *Wiley Interdisciplinary Reviews: Computational Statistics, 4(3),* 275–294. https://doi.org/10.1002/wics.1198

8. Zeng, L., Wu, W., & Li, Z. (2023). Research on multiple regression linear prediction model of 3D printing process parameters. In X. Yuan & G. Wu (Eds.), *Third International Conference on Mechanical, Electronics, and Electrical and Automation Control (METMS 2023)* (Vol. 12722, p. 127224Q). SPIE. https://doi.org/10.1117/12.2679759

9. Khan, M. F., Alam, A., Siddiqui, M. A., Alam, M. S., Rafat, Y., Salik, N., & Al-Saidan, I. (2021). Real-time defect detection in 3D printing using machine learning. *Materials Today: Proceedings, 42,* 521–528. https://doi.org/10.1016/j.matpr.2020.10.482

10. Goh, G. D., & Yeong, W. Y. (2021). Applications of machine learning in 3D printing. *Materials Today: Proceedings, 42,* 521–528. https://doi.org/10.1016/j.matpr.2020.10.482

11. Malik, A. W., Mahmood, M. A., & Liou, F. (2024). Digital twin–driven optimization of laser powder bed fusion processes: A focus on lack-of-fusion defects. *Rapid Prototyping Journal*. https://doi.org/10.1108/rpj-02-2024-0091

12. Sampedro, G. A. R., Putra, M. A. P., & Abisado, M. (2023). 3D-AmplifIER: An ensemble machine learning approach to digital twin fault monitoring for additive manufacturing in smart factories. *IEEE Access, 11,* 64128–64140. https://doi.org/10.1109/access.2023.3289536

13. Kantaros, A., Piromalis, D., Tsaramirsis, G., Papageorgas, P., & Tamimi, H. (2021). 3D printing and implementation of digital twins: Current trends and limitations. *Applied System Innovation, 5(1),* 7. https://doi.org/10.3390/asi5010007

14. Andreu, A., Kim, S., Kim, I., Kim, J.-H., Noh, J., Lee, S., Lee, W., Su, P.-C., & Yoon, Y.-J. (2024). Processing Challenges and Delamination Prevention Methods in Titanium-Steel DED 3D Printing. *International Journal of Precision Engineering and Manufacturing Green Technology*. https://doi.org/10.1007/s40684-024-00598-9

15. Duda, T., & Raghavan, L. V. (2016). Durrës. *3D Metal Printing Technology*, *49*(29), 103–110. https://doi.org/10.1016/j.ifacol.2016.11.111

16. Svetlizky, D., Zheng, B., Andrey Vyatskikh, Das, M., Bose, S., Bandyopadhyay, A., Schoenung, J. M., Lavernia, E. J., & Noam Eliaz. (2022). Laser-based directed energy deposition (DED-LB) of advanced materials. *Materials Science and Engineering A-Structural Materials Properties Microstructure and Processing*, *840*, 142967–142967. https://doi.org/10.1016/j.msea.2022.142967

17. Choi, S.-W., Kim, Y.-S., Yum, Y.-J., & Yang, S.-Y. (2020). A Study on Strengthening Mechanical Properties of a Punch Mold for Cutting by Using an HWS Powder Material and a DED Semi-AM Method of Metal 3D Printing. *Journal of Manufacturing and Materials Processing*, *4*(4), 98. https://doi.org/10.3390/jmmp4040098

18. Zoubeir, A., & Radi, B. (2024). Finite Element Modeling (FEM) of the thermal behavior of 3D printed parts during Directed Energy Deposition (DED). *Incertitudes Et Fiabilité Des Systèmes Multiphysiques*, *8*(1), 36–39. https://doi.org/10.21494/iste.op.2024.1223

19. *Modeling of the selective laser melting (SLM) process using the ANSYS software - ProQuest*.(n.d.).https://www.proquest.com/docview/2768680035?pq-orgsite=gscholar&fromopenview=true&sourcetype=Dissertations%20&%20Theses

20. Maurya, J. C., Nagarjuna, A. S. V., & Mallaiah, M. (2024). Three-Dimensional Numerical Simulation of Laser Metal Deposition Method with Bead Profile Optimization. In *Lecture notes in mechanical engineering* (pp. 263–280). https://doi.org/10.1007/978-981-97-6016-9_22