

MPGT: Multimodal physics-constrained graph transformer learning for hybrid digital twins

1st Benjamin Uhrich

Leipzig University

*Center for Scalable Data Analytics
and Artificial Intelligence (ScaDS.AI)*

Dresden/Leipzig, Germany

uhrich@informatik.uni-leipzig.de

2nd Erhard Rahm

Leipzig University

*Center for Scalable Data Analytics
and Artificial Intelligence (ScaDS.AI)*

Dresden/Leipzig, Germany

rahm@informatik.uni-leipzig.de

Abstract—In additive manufacturing (AM), heating and cooling cycles are the main factors influencing component quality. Physics-based models are the most powerful tool for simulating such processes. However, the increasing availability of multimodal sensor data in complex industrial systems requires hybrid digital twins. Machine learning and physics-based process simulation capabilities are crucial to optimize the complex production environment. To integrate and analyze multimodal sensor data while capturing the underlying physical behaviour, we propose a multimodal physics-constrained graph transformer learning framework (MPGT). The framework is made up of three parts: A graph neural diffusion approach is employed to model the continuous heat transport in AM and predict temperature features. An autoencoder and principal component analysis (PCA) generate dynamic feature embeddings representing potential structural weaknesses, such as corners and edges on the top surface of the printed components. The intermediate fused features are processed by a temporal-spatial graph transformer, which aggregates multimodal information to predict the component quality. Our method enables precise defect identification and predictive insight into component integrity. Experimental evaluations of grayscale images and thermal data demonstrate the effectiveness of our framework in accurately predicting quality outcomes for 3D-printed components. This emphasises the potential of the framework for complex industrial applications in engineering and manufacturing. The MPGT code will be made open-sourced on GitHub once the paper is published.

Index Terms—multimodal machine learning, decision making, graph transformer, heat transfer, additive manufacturing

I. INTRODUCTION

Additive manufacturing (AM) has emerged as a transformative technology in modern engineering, enabling the production of highly customized components with complex geometries. However, ensuring the quality of these components remains a significant challenge due to the intricate thermal dynamics involved during the manufacturing process. Among these dynamics, heating and cooling cycles are identified as critical factors influencing the structural integrity and overall quality of the printed components. Effective condition monitoring and quality prediction are essential to minimize defects and optimize production outcomes in AM.

The increasing availability of multimodal sensor data in complex engineering systems presents new opportunities and challenges for comprehensive analysis of manufacturing pro-

cesses. In particular, grayscale images documenting surface characteristics and thermal data capturing heat transport dynamics offer complementary perspectives on the quality-related phenomena in 3D printing. However, analyzing and integrating these diverse data modalities to extract actionable insights requires intelligent approaches that combine domain-specific physical knowledge with advanced data learning capabilities. To address this challenge, we propose a multimodal physics-constrained graph transformer model (MPGT) for hybrid digital twins that enables condition monitoring and quality prediction of 3D-printed components. Our framework leverages a multimodal dataset comprising grayscale images and thermal data. Central to our approach is graph neural diffusion, enhanced by physics constraints to model temperature distribution and learn temperature-based features. The employment of an autoencoder and principal component analysis (PCA) facilitates the generation of dynamic feature embeddings of the grayscale images, thus representing potential structural weaknesses.

The intermediate fused features are processed by a graph transformer, which aggregates multimodal information across space and time to deliver a robust quality prediction. By combining domain knowledge of heat transport dynamics with graph learning, our approach achieves precise identification of potential defects and provides predictive insights into component integrity. Experimental evaluations demonstrate the effectiveness of our framework, highlighting its potential for complex industrial applications in engineering and manufacturing.

- The proposed approach combines an autoencoder with physics-constrained graph neural diffusion and a temporal-spatial graph transformer for condition monitoring and quality prediction of 3d printed components.
- We use dynamically modified latent space of an autoencoder and PCA to create feature embeddings of grayscale images.
- We evaluate our framework on a real measurement mul-

timodal data set of grayscale images and thermal data.

The remainder of this paper is organized as follows: section II reviews related work in multimodal machine learning, graph transformer networks and quality prediction in AM. section III introduces the architecture and methodological details of MPGT. section IV outlines the experimental setup and presents results demonstrating the efficacy of our approach. section V discusses the implications of the findings and concludes the paper with future research directions.

II. RELATED WORK

Some of the preliminary work done to lay the groundwork for our work is described below.

A. 3D Printing

The advent of AM has profoundly influenced the landscape of manufacturing and engineering. Laser Powder Bed Fusion (LPBF) is a key AM technology used for creating complex and high-precision metal components. In LPBF, a heat source (either a laser or an electric beam) selectively melts powder particles in a layer-by-layer manner, thereby building a three-dimensional object based on a digital model. One of the primary challenges in LPBF is the reduction and avoidance of defects, such as porosity, surface roughness, and thermal stresses, which can lead to enhanced structural component quality [1]–[3]. Du *et al.* investigate pore defects in LPBF and their formation mechanism [4]. Wang *et al.* employed an overview of melt pool characteristics for process optimization in LPBF [5]. Mukherjee *et al.* developed a model to calculate heat transfer and fluid flow, with the aim of determining temperature and velocity fields for a wide variety of materials [6], [7]. In the present work, our investigation focuses on three-dimensional printed components of 316L stainless steel.

B. Physics-constrained Machine Learning

In recent years, machine learning has experienced a significant increase in complex technical systems, largely driven by the proliferation of sensor data. However, despite the growing availability of sensing and data in general, there remains a significant challenge in fully characterizing numerous engineering systems and structures, such as 3D printing, through a purely data-driven approach. This underscores the necessity for physical knowledge to complement sensor data, thereby necessitating the development of physics-constrained machine learning methodologies. The incorporation of physical domain knowledge, most typically expressed through partial differential equations (PDEs), into the loss function is a critical step in ensuring the preservation of underlying physical laws. This integration can facilitate the handling of limited data. Raissi *et al.* initially developed a physics-informed neural network and subsequently demonstrated the efficacy of this approach in solving both forward and inverse problems involving nonlinear partial differential equations [8]. Cai *et al.* presented an overview of physics-informed neural networks for fluid mechanics [9]. Wessels *et al.* developed the neural particle method for computational fluid dynamics

[10]. Cai *et al.* presented physics-informed neural networks for heat transfer problems [9]. Hu *et al.* addressed the curse of dimensionality with physics-informed neural networks [11]. Uhrich *et al.* employed physics-informed deep learning to predict heat transfer and quantify anomalies for real-time fault mitigation in 3D printing [12] [13]. In addition, Uhrich *et al.* developed a neural network inspired by a physical model of an electrodynamic valve system for the prediction of valve failures [14]. Our work employs the concept of physics-constrained machine learning, integrating physical principles of heat transfer into the loss function.

C. Graph Transformer Networks

Graph transformer networks (GTNs) represent a significant evolution in graph-based deep learning, combining the strengths of graph neural networks (GNNs) and transformer with self attention mechanism [15], [16] [17]. In recent years, these methods have gained prominence due to their ability to leverage global attention mechanisms and structural encodings to capture both local and global dependencies in graph data effectively [18]. Liu *et al.* developed a spatial-temporal multigraph transformer network for joint prediction of multiple vessel trajectories [19]. Chen *et al.* presented a survey on graph neural networks and graph transformers in computer vision [20]. Geng *et al.* developed a dynamic-learning spatial-temporal transformer network for accurate vehicular trajectory prediction at urban intersections, addressing challenges in modeling temporal dependencies and spatial interactions [21]. Feng *et al.* have developed a novel energy-informed graph transformer that has demonstrated its potential to improve the efficiency, robustness and accuracy of network approximation methods for both forward and inverse problems in the field of solid mechanics [22]. Wang *et al.* proposed a novel anomaly detection and classification architecture for multivariate time series data in industrial and IoT systems. This architecture combines a graph learning method and a high-efficiency transformer model to perform rapid and accurate sequence modeling [23]. In the present study, the integration of a temporal-spatial graph transformer layer facilitates the aggregation of temperature features and top surface patterns, thereby enabling the prediction of component quality. This approach enables multimodal machine learning.

D. Multimodal Machine Learning

Multimodal machine learning focuses on integrating and analyzing data from multiple modalities, such as text, images, audio, video, and sensor data, to enable a more comprehensive understanding of complex systems. This approach leverages the complementary information across modalities, addressing challenges such as modality alignment, fusion, and co-dependencies. Baltrušaitis *et al.* presented a survey and taxonomy of multimodal machine learning [24]. Liang *et al.* presented principles, challenges and open questions of multimodal machine learning [25]. A review of multimodal machine learning approaches in healthcare are presented by Kroner *et al.* [26]. Lee developed a multimodal machine

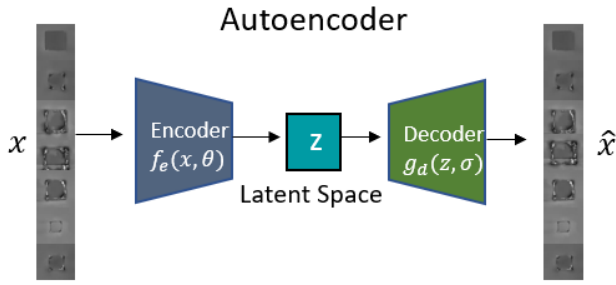


Fig. 1. **Autoencoder architecture** - The purpose of the encoder is to transform input data into latent space, with the objective of reconstructing the input data using a decoder

learning model for predicting heat transfer characteristics in micro-pin fin heat sinks [27]. Bauer *et al.* investigated the correlation of multimodal data in additive manufacturing [28]. Song *et al.* presented a comprehensive overview of the current state, advancements and challenges of multimodal machine learning in regard of engineering design [29].

III. MULTIMODAL PHYSICS-CONSTRAINED GRAPH TRANSFORMER NETWORK

In this section the functionality of all learning components of our MPGT are introduced.

A. Autoencoder

An autoencoder is composed of two neural networks: an encoder and a decoder (see Figure 1). The encoder is trained to identify a transformation function

$$z = f_e(x, \theta) \quad (1)$$

that maps the input feature space to a low-dimensional vector space. The latent space generated by the encoder is then transformed by the learned decoder,

$$\hat{x} = g_d(z, \sigma) \quad (2)$$

resulting in the reconstruction of the original feature space [30]. In order to advance the learning of the encoder-decoder function based on Equation 1, Equation 2

$$\hat{x} = g_d(f_e(x, \theta), \sigma) \quad (3)$$

it is required to minimize the loss function:

$$\mathcal{L}_{auto} = \frac{1}{N_a} \sum_{i=1}^{N_a} (\hat{x}_i - x_i)^2 \quad (4)$$

Application tasks of autoencoders in general are dimension reduction and feature extraction [31].

B. Physics-constrained Graph Neural Diffusion

1) *Graph Neural Diffusion*: Chamberlain *et al.* have determined that the process by which information is propagated through the individual layers of a GNN can be described in a manner analogous to the calculation of the approximate solution of a continuous diffusion process. This interpretation

of GNN layers as discrete approximations of an underlying PDE is referred to as graph neural diffusion [32]. In order to predict the heat transfer process within the entire component using temperature values derived from thermal sensor data of the top surface, we propose the introduction of our graph neural diffusion model based on the explicit Euler method and the heat equation. The idea was first described in [33]:

$$T_{n+1}(T_n) = T_n + \delta t \hat{L}(T_n)T_n + Q(T_n) \quad (5)$$

with the Temperature state T_n , the laplacian matrix \hat{L} as discrete version of the laplace operator allowing temperature dependent conductivity and a dissipation vector Q to account for heat loss at the boundaries. The Laplacian matrix can be described as:

$$\hat{L} = \hat{D} - \hat{A} \quad (6)$$

with adjacency matrix \hat{A} and diagonal matrix \hat{D} . The state-dependent adjacency matrix \hat{A} is defined as:

$$\hat{A} = \begin{cases} 0, & a_{ij} = 0, \\ c_{ij}, & a_{ij} \neq 0 \end{cases}, \quad (7)$$

where the non-zero entries c_{ij} of \hat{A} are estimated by a learnable function

$$c_{ij} = \varphi(\varrho_{ij}, T_n(v_i), T_n(v_j), C(v_i), C(v_j), d(v_i), d(v_j)) \quad (8)$$

In a similar fashion the dissipation vector can be learned:

$$Q_i(T_n) = \psi(T_n(v_i), C(v_i), d(v_i)) \quad (9)$$

In order to build the functional dependencies of the temperature state, the local graph structure and the local temperature values influence the entries of L and Q . The local information for a single vertex v_i is given by the temperature $T_n(v_i)$, the spatial property of the vertex $C(v_i)$, and the scale-invariant density $d(v_i)$. For an edge between two adjacent vertices v_i and v_j , the local information is defined as the distance between the vertex $\varrho_{ij} = \|v_i - v_j\|_2$, together with the local information for each of the two vertices. The spatial property of the vertex is defined by vertex classes (top boundary, side boundary, bottom boundary and interior). The scale invariant density is defined as:

$$d(v_i) = \left(\frac{8}{3}\pi\right)^{-1} \sum_{j \neq i} \|v_i - v_j\|_2^{-2} \quad (10)$$

2) *Physical Constraints*: In order to accelerate the learning process and improving the prediction accuracy, physical principles of heat transport and mathematical conclusions can be integrated into the loss function. It is known from the discretization of the continuous Laplacian that the connectivity of two adjacent vertices should be proportional to the inverse square of their distance:

$$\mathcal{L}_\varphi = \sum_{i,j:i \sim j} \left(\varphi(\varrho_{ij}) - \frac{1}{\varrho_{ij}^2} \right) \quad (11)$$

Dissipation can only occur at the boundaries, for points of the interior class, $\psi = 0$ is required for the learned entries of our dissipation vector:

$$\mathcal{L}_\psi = \sum_{i: C(v_i)=\text{int.}} \psi(T_n(v_i), C(v_i), d(v_i))^2 \quad (12)$$

Using knowledge from the theoretical study of PDEs, statements about the temporal evolution of the heat states can be made. The total thermal energy in the body can only change because of dissipation:

$$\mathcal{L}_{\text{heat}} = \left(\sum_i \left(T_{n+1}(v_i) - T_n(v_i) + Q_n(v_i) \right) \right)^2. \quad (13)$$

Another well known property of the evolution of heat distribution is the maximum principle:

$$\mathcal{L}_{\text{max}} = \sum_i \max \left(0, T_{n+1}(v_i) - \max(M) \right)^2 \quad (14)$$

$$\mathcal{L}_{\text{min}} = \sum_i \max \left(0, \min(M) - T_{n+1}(v_i) \right)^2 \quad (15)$$

$$M = \{T_n(v_i); T_{n+1}(v_j), v_j \sim v_i\}$$

The temperature at a vertex is within the range given by the minimum and maximum over its previous temperature and the temperatures of connected vertices. The discrete approximation of the potential energy can be defined as:

$$\mathcal{L}_{\text{energy}} = \max \left(0, \widehat{E}(T_{n+1}) - \widehat{E}(T_n) \right) \quad (16)$$

where

$$\widehat{E}(T_n) = \frac{1}{N} \sum_i \left(T_n(v_i) - \bar{T}_n \right)^2$$

$$\text{with } \bar{T}_n = \frac{1}{N} \sum_i T_n(v_i)$$

The potential energy is higher when temperature differences are large. The loss ensures that the discrete energy representing temperature variance, decreases over time to respect physical laws of dissipation. It penalizes any increase in energy between consecutive time steps. Finally the best fit of the real measurement data on the top surface and the prediction is needed:

$$\mathcal{L}_{\text{Tdata}} = \sum_{i: C_i=\text{top}} \left(T_{n+1}(v_i) - T_{n+1}^{(\text{data})}(v_i) \right)^2 \quad (17)$$

For training the model, the regularization loss functions as well as the prediction loss are added using corresponding weights.

C. Temporal-spatial Graph Transformer Layer

As described in subsection II-C the temporal-spatial graph transformer network is designed to process graph structured multimodal features. Our architecture combines graph attention mechanisms and deep learning to learn node and graph-level representations effectively. Each node v_i has an input feature vector $x_i \in \mathbb{R}^{C_{in}}$, with the number of input features C_{in} (e.g., 2 for processed temperature features and grayscale

embedding). The input features of N nodes are transformed into a shared hidden space:

$$H = (h_0, h_1, h_2, \dots, h_N) \quad (18)$$

where

$$h_i = W_e x_i + b_e \quad (19)$$

The graph transformer layers aggregate and transform node features using multi-head attention. Let H represent the node embeddings. The attention mechanism is defined as:

$$\alpha_{ij} = \text{softmax}_j \left(\frac{qk^T}{\sqrt{d}} \right) \quad (20)$$

$$= \text{softmax}_j \left(\frac{W_q h_i (W_k h_j)^T}{\sqrt{d}} \right)$$

with learnable weight matrices W_q, W_k . q and k are key and query vectors. The node embeddings are updated using:

$$h_i = \sum_{j \in \mathcal{N}(i)} \alpha_{ij} W_v h_j + W_c c_{ij} \quad (21)$$

with learnable weight matrices W_v and W_c . c_{ij} are the edge attributes from our learned connectivity model Equation 8. To derive a graph-level representation, global mean pooling aggregates node-level embeddings:

$$h_G = \frac{1}{|V|} \sum_{i \in V} h_i \quad (22)$$

with the nodes set V in the graph. The optimisation of the weights is accelerated by the loss function that measures the prediction quality (categorical cross-entropy loss) on the training data:

$$\mathcal{L}_{\text{cross}}(f) = \frac{1}{|\mathcal{D}|} \sum_{(x,y) \in \mathcal{D}} \sum_{j=1} -\ln(f(x)_j) y_j \quad (23)$$

IV. RESULTS

A. Experimental data

In the printing machines, two cameras are integrated to capture images during the printing process. One grayscale image with a resolution of 382×288 pixels is produced for each completed print layer. Additionally, thermal images, recorded in the wavelength range of $\lambda = 7\mu\text{m}$ to $14\mu\text{m}$, are captured at a frequency of 3 Hz throughout the entire printing process.

In order to train the proposed MPGT, we utilize data from two components: One that maintains high quality throughout the printing process and another that starts with high quality but becomes deformed as printing progresses. The dataset comprises three distinct classes of components: high quality, deficient (warning), and low quality. Grayscale images representing the surface pattern of each completed print layer, along with thermal images documenting the heat transport behavior on the top surface, are used as input.

The thermal images are transformed into graph-structured data $G = (V, E)$, where $V = 1, \dots, N$ represents the set of vertices (nodes), and $E \subseteq V \times V$ represents the set of

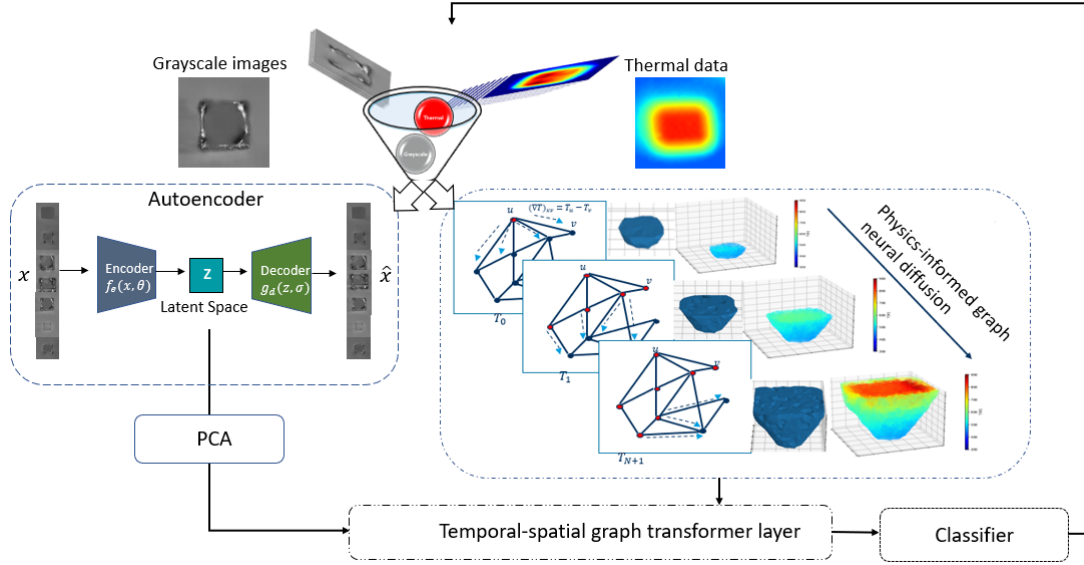


Fig. 2. **MPGT Model Architecture** - Flowchart of data processing, feature prediction and visualization

edges. This graph structure serves as a representation of the physical object, specifically the 3D-printed component. For nodes corresponding to the top boundary class, temperature values are directly assigned based on real measurement data. For interior and side boundary nodes, temperature features are predicted using the proposed physics-informed graph neural diffusion method.

The incorporation of grayscale images into the graph-structured data is achieved by utilising the proposed autoencoder and PCA to generate a dynamically changing latent feature space. This dynamic latent feature space is necessary because the surface size of the component changes during the printing process. By integrating the features from both modalities, MPGT is able to model the thermal and structural behaviors of the component more effectively.

B. Intermediate Fusion

The feature embedding generated by the autoencoder and the predicted temperature distribution are fused to create the input feature vector for the graph transformer. In order to train the MPGT to learn the internal heat dynamics, recognize anomalies, and capture the influence of the printed shape, all proposed loss functions Equation 4, Equation 11, Equation 12, Equation 13, Equation III-B2, Equation 16, Equation 17 are combined in the training objective:

$$\begin{aligned} \mathcal{L}_{MPGT} = & \lambda_1 \mathcal{L}_{auto} \\ & + \lambda_2 \mathcal{L}_\varphi \\ & + \lambda_3 \mathcal{L}_\psi \\ & + \lambda_4 \mathcal{L}_{heat} \\ & + \lambda_5 \mathcal{L}_{energy} \\ & + \lambda_6 \mathcal{L}_{Tdata} \\ & + \lambda_7 \mathcal{L}_{cross} \end{aligned} \quad (24)$$

TABLE I
PERFORMANCE OF DIFFERENT CLASSIFICATION MODELS - EVALUATION WERE DONE USING BALANCED ACCURACY (ACC.) AND f_1 SCORES FOR SUCCESSFUL (S), WARNING (W), AND FAILED (F).

Model	Acc.	$f_{1,s}$	$f_{1,w}$	$f_{1,f}$
Resnet18	0.917	0.98	0.87	0.89
Resnet50	0.920	0.98	0.87	0.90
EfficientNet-B0	0.926	0.98	0.88	0.89
MobileNet-V3	0.860	0.96	0.79	0.81
MPGT	0.954	0.98	0.91	0.95

This integrated approach ensures that the model effectively leverages multimodal data to make accurate predictions about the quality and behavior of the 3D-printed components. The model architecture and the flowchart of the process can be seen in Figure 2.

The model is trained on 13,944 timepoints, corresponding to a printing duration of 77.47 minutes, and evaluated on 18,981 timepoints. Among these, 7,892 timepoints correspond to high-quality states, 5,299 represent anomalies that reduce component quality (such as corners and edges), and 5,790 timepoints indicate low-quality states where the structure is already deformed. Table I shows the performance of the model, achieving a prediction accuracy of 95.4%. Additionally, the F1-scores for the three quality classes are calculated to provide a more comprehensive evaluation of the model's performance. The MPGT approach is compared against four other neural network architectures, which were used in a previous study to investigate and predict component quality [28]. The proposed MPGT model is distinct from alternative models in that it is trained on a combination of multimodal features and the temporal evolution of heat states. In contrast, the other models are trained exclusively on temperature features of a thermal image of one print layer and can thus only predict the quality

of a finished print layer.

V. DISCUSSION AND CONCLUSION

In this paper, we presented a physics-constrained graph transformer framework for analyzing multimodal sensing and processing in engineering. The component to be manufactured was modelled as a graph, thereby enabling the integration and analysis of both thermal data and structural characteristics through a combination of advanced sub-learning models. Specifically, an autoencoder and physics-informed graph neural diffusion were employed to predict node features for the graph. These features were aggregated using a temporal-spatial graph transformer, which facilitated decision-making about component quality. Furthermore, the learned edge weights from the graph neural diffusion were utilized as edge attributes in the graph transformer layer, enhancing its interpretability and accuracy.

Our proposed intelligent model demonstrates the ability to understand underlying heat transfer processes and their effects on physical structures. A unique aspect of our approach is its capability for time-state prediction, effectively functioning as a digital heat twin to simulate and predict heat transport behavior. Additionally, the autoencoder predicts areas of potential structural weaknesses, enabling proactive monitoring and quality assessment.

This modular framework offers versatility: its components (e.g., the digital heat twin and structural anomaly prediction) can function independently or synergistically. Together, they provide a comprehensive system for monitoring the condition of components and guiding the progress of the overall manufacturing process. It has been demonstrated that the integration and learning from multiple modalities of sensing can yield benefits, including enhanced prediction accuracy when compared with the utilisation of temperature features alone. Furthermore, the model under consideration is capable of predicting temporal states within a print layer, thereby facilitating sensitive monitoring. Future work could explore scaling the framework to more complex geometries and also provide advice on what to do if the component gets too hot during printing, to make sure the quality of future print layers is maintained.

REFERENCES

- [1] S. Chowdhury, N. Yadaiah, C. Prakash, S. Ramakrishna, S. Dixit, L. R. Gupta, and D. Buddhi, "Laser powder bed fusion: a state-of-the-art review of the technology, materials, properties & defects, and numerical modelling," *Journal of Materials Research and Technology*, vol. 20, pp. 2109–2172, 2022.
- [2] T. DebRoy, H. L. Wei, J. S. Zuback, T. Mukherjee, J. W. Elmer, J. O. Milewski, A. M. Beese, A. d. Wilson-Heid, A. De, and W. Zhang, "Additive manufacturing of metallic components—process, structure and properties," *Progress in materials science*, vol. 92, pp. 112–224, 2018.
- [3] H. D. Nguyen, A. Pramanik, A. Basak, Y. Dong, C. Prakash, S. Debnath, S. Shankar, I. Jawahir, S. Dixit, and D. Buddhi, "A critical review on additive manufacturing of ti-6al-4v alloy: Microstructure and mechanical properties," *Journal of Materials Research and Technology*, vol. 18, pp. 4641–4661, 2022.
- [4] C. Du, Y. Zhao, J. Jiang, Q. Wang, H. Wang, N. Li, and J. Sun, "Pore defects in laser powder bed fusion: Formation mechanism, control method, and perspectives," *Journal of Alloys and Compounds*, vol. 944, p. 169215, 2023.
- [5] J. Wang, R. Zhu, Y. Liu, and L. Zhang, "Understanding melt pool characteristics in laser powder bed fusion: An overview of single- and multi-track melt pools for process optimization," *Advanced Powder Materials*, vol. 2, no. 4, p. 100137, 2023.
- [6] T. Mukherjee, H. Wei, A. De, and T. DebRoy, "Heat and fluid flow in additive manufacturing—part i: Modeling of powder bed fusion," *Computational Materials Science*, vol. 150, pp. 304–313, 2018.
- [7] —, "Heat and fluid flow in additive manufacturing—part ii: Powder bed fusion of stainless steel, and titanium, nickel and aluminum base alloys," *Computational Materials Science*, vol. 150, pp. 369–380, 2018.
- [8] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational physics*, vol. 378, pp. 686–707, 2019.
- [9] S. Cai, Z. Mao, Z. Wang, M. Yin, and G. E. Karniadakis, "Physics-informed neural networks (pinns) for fluid mechanics: A review," *Acta Mechanica Sinica*, vol. 37, no. 12, pp. 1727–1738, 2021.
- [10] H. Wessels, C. Weißenfels, and P. Wriggers, "The neural particle method—an updated lagrangian physics informed neural network for computational fluid dynamics," *Computer Methods in Applied Mechanics and Engineering*, vol. 368, p. 113127, 2020.
- [11] Z. Hu, K. Shukla, G. E. Karniadakis, and K. Kawaguchi, "Tackling the curse of dimensionality with physics-informed neural networks," *Neural Networks*, vol. 176, p. 106369, 2024.
- [12] B. Uhrich, M. Schäfer, O. Theile, and E. Rahm, "Using physics-informed machine learning to optimize 3d printing processes," in *International Conference of Progress in Digital and Physical Manufacturing*. Springer, 2021, pp. 206–221.
- [13] B. Uhrich, N. Pfeifer, M. Schäfer, O. Theile, and E. Rahm, "Physics-informed deep learning to quantify anomalies for real-time fault mitigation in 3d printing," *Applied Intelligence*, vol. 54, no. 6, pp. 4736–4755, 2024.
- [14] B. Uhrich, N. Hlubek, T. Häntschel, and E. Rahm, "Using differential equation inspired machine learning for valve faults prediction," in *2023 IEEE 21st International Conference on Industrial Informatics (INDIN)*. IEEE, 2023, pp. 1–8.
- [15] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.
- [16] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio et al., "Graph attention networks," *stat*, vol. 1050, no. 20, pp. 10–48 550, 2017.
- [17] A. Vaswani, "Attention is all you need," *Advances in Neural Information Processing Systems*, 2017.
- [18] S. Yun, M. Jeong, R. Kim, J. Kang, and H. J. Kim, "Graph transformer networks," *Advances in neural information processing systems*, vol. 32, 2019.
- [19] R. W. Liu, W. Zheng, and M. Liang, "Spatio-temporal multi-graph transformer network for joint prediction of multiple vessel trajectories," *Engineering Applications of Artificial Intelligence*, vol. 129, p. 107625, 2024.
- [20] C. Chen, Y. Wu, Q. Dai, H.-Y. Zhou, M. Xu, S. Yang, X. Han, and Y. Yu, "A survey on graph neural networks and graph transformers in computer vision: A task-oriented perspective," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.
- [21] M. Geng, Y. Chen, Y. Xia, and X. M. Chen, "Dynamic-learning spatial-temporal transformer network for vehicular trajectory prediction at urban intersections," *Transportation research part C: emerging technologies*, vol. 156, p. 104330, 2023.
- [22] B. Feng and X. Zhou, "Energy-informed graph transformer model for solid mechanical analyses," *Communications in Nonlinear Science and Numerical Simulation*, p. 108103, 2024.
- [23] C. Wang and G. Liu, "From anomaly detection to classification with graph attention and transformer for multivariate time series," *Advanced Engineering Informatics*, vol. 60, p. 102357, 2024.
- [24] T. Baltrušaitis, C. Ahuja, and L.-P. Morency, "Multimodal machine learning: A survey and taxonomy," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 2, pp. 423–443, 2018.
- [25] P. P. Liang, A. Zadeh, and L.-P. Morency, "Foundations & trends in multimodal machine learning: Principles, challenges, and open questions," *ACM Computing Surveys*, vol. 56, no. 10, pp. 1–42, 2024.
- [26] F. Krones, U. Marikkar, G. Parsons, A. Szmul, and A. Mahdi, "Review of multimodal machine learning approaches in healthcare," *Information Fusion*, vol. 114, p. 102690, 2025.

- [27] H. Lee, G. Lee, K. Kim, D. Kong, and H. Lee, "Multimodal machine learning for predicting heat transfer characteristics in micro-pin fin heat sinks," *Case Studies in Thermal Engineering*, vol. 57, p. 104331, 2024.
- [28] M. Bauer, B. Uhrich, M. Schäfer, O. Theile, C. Augenstein, and E. Rahm, "Multi-modal artificial intelligence in additive manufacturing: Combining thermal and camera images for 3d-print quality monitoring." in *ICEIS (1)*, 2023, pp. 539–546.
- [29] B. Song, R. Zhou, and F. Ahmed, "Multi-modal machine learning in engineering design: A review and future directions," *Journal of Computing and Information Science in Engineering*, vol. 24, no. 1, p. 010801, 2024.
- [30] J. Zhai, S. Zhang, J. Chen, and Q. He, "Autoencoder and its various variants," in *2018 IEEE international conference on systems, man, and cybernetics (SMC)*. IEEE, 2018, pp. 415–419.
- [31] S. Phadikar, N. Sinha, and R. Ghosh, "Unsupervised feature extraction with autoencoders for eeg based multiclass motor imagery bci," *Expert Systems with Applications*, vol. 213, p. 118901, 2023.
- [32] B. Chamberlain, J. Rowbottom, M. I. Gorinova, M. Bronstein, S. Webb, and E. Rossi, "Grand: Graph neural diffusion," in *International conference on machine learning*. PMLR, 2021, pp. 1407–1418.
- [33] B. Uhrich, T. Häntschel, M. Schäfer, and E. Rahm, "Neural diffusion graph convolutional network for predicting heat transfer in selective laser melting," in *International Workshop on Combinatorial Image Analysis*. Springer, 2024, pp. 150–164.