Progress in Additive Manufacturing
https://doi.org/10.1007/s40964-026-01854-x

RESEARCH

Learning composition-sensitive signatures in multi-material PBF-LB: a lightweight, modality-aware, explainable graph-attention sensor fusion framework for in-situ monitoring of graded 316L–CuCrZr alloys

Vigneashwara Pandiyan[1] · **Antonios Baganis**[2] · **Antti Salminen**[1] · **Christian Leinenbach**[2]

Received: 21 January 2026 / Accepted: 6 July 2026

Abstract
Decoding composition-sensitive process signatures in Laser Powder Bed Fusion (PBF-LB) of multi-material structures remains challenging due to the complex and transient nature of melt-pool emissions across spatially varying material gradients. In this study, we propose a modality-aware learning framework that fuses acoustic emission (AE) and back-reflected optical emission (OE) at the laser wavelength for spatiotemporally resolved classification of local composition in graded 316L–CuCrZr alloys. The framework integrates learnable shapelet extraction with graph-based attention, where shapelets act as interpretable descriptors of modality-specific temporal behavior and enable compact, discriminative representation learning. The resulting dual-modality representations are organized as a temporal graph of signal segments, and a Graph Attention Network (GAT) with modality-specific attention heads adaptively prioritizes sensor streams based on their relevance to compositional variation. Across five Cu-containing concentrations (20%–100 wt.%), the model achieved up to 92% accuracy with only ~4,000 trainable parameters. Beyond accuracy, the results show that sensor fusion is imperative for multi-material PBF-LB, as no single modality consistently captures composition-dependent phenomena across the CuCrZr gradient. The learned attention and saliency trends further reveal that OE becomes increasingly influential at higher CuCrZr fractions, consistent with increased reflectivity at the laser wavelength. Overall, the proposed framework establishes a compact and interpretable approach for decoding composition-dependent multimodal process signatures in multi-material PBF-LB.

Keywords Laser powder bed fusion · Multi-material additive manufacturing · Multi-material process monitoring · Acoustic emission · Optical emission · Graph networks

1 Introduction

Metal additive manufacturing (AM) has emerged as a transformative manufacturing paradigm, enabling the direct fabrication of complex, high-performance components with unprecedented geometric freedom and material efficiency [1–3]. Unlike conventional subtractive or formative techniques, AM builds parts layer by layer from digital models, reducing material waste, shortening lead times, and facilitating design customization. These advantages have spurred significant industrial adoption across aerospace [4], automotive [5], energy [6], and biomedical sectors [7], where the ability to produce lightweight, topology-optimized, and functionally graded structures offers substantial performance gains and cost benefits. Despite these advantages, the widespread deployment of metal AM remains constrained

✉ Vigneashwara Pandiyan
vigneashwara.solairajapandiyan@utu.fi

Antonios Baganis
antonios.baganis@polytechnique.edu

Antti Salminen
antti.salminen@utu.fi

Christian Leinenbach
Christian.Leinenbach@empa.ch

1 University of Turku, Turku, Finland

2 Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland

Published online: 23 July 2026

by the challenges associated with part qualification and certification [8]. In many high-value industries, the stringent performance and safety requirements demand exhaustive post-process inspection, including Non-Destructive Evaluation and mechanical testing, to ensure component integrity [9]. These procedures are often time-consuming, costly, and incompatible with the on-demand, agile manufacturing workflows that AM is intended to enable [10, 11]. Moreover, the stochastic nature of melt pool dynamics, powder spreading, and layer-to-layer thermal interactions introduces variability that complicates process control and makes conventional trial-and-error parameter tuning inefficient [12, 13]. A promising pathway to overcoming these barriers is the integration of real-time process monitoring and control into AM workflows [14]. By capturing high-fidelity, in-situ sensor data during fabrication, it becomes possible to detect process instabilities, identify defect precursors, and infer material properties without reliance on exhaustive downstream inspection [15, 16]. Such process-embedded intelligence enables the development of closed-loop control strategies, accelerates part qualification, and opens the door to adaptive manufacturing, where process parameters can be dynamically tuned in response to evolving build conditions [17, 18]. This shift from post-build inspection to in-process monitoring represents a critical step toward scalable, cost-effective adoption of metal AM in safety–critical industries [19].

Laser Powder Bed Fusion has established itself as a leading metal AM technique, capable of producing fully dense metallic components with exceptional geometric accuracy [20, 21]. The process employs a tightly focused laser to selectively melt successive layers of metal powder, guided directly by digital design data. This layer-wise fabrication enables the creation of intricate features, internal channels, and complex lattice architectures that are often impractical or impossible to achieve using conventional manufacturing methods. Laser Powder Bed Fusion routinely delivers near-theoretical part densities, with mechanical properties comparable to or exceeding those of wrought counterparts, making it highly attractive for demanding applications [22]. Despite these advantages, PBF-LB is inherently sensitive to fluctuations in parameters such as laser power, scan speed, and layer deposition quality, owing to the localized nature of the melt pool and the rapid thermal cycles involved [23]. Slight perturbations in process conditions can result in defects including porosity, lack of fusion, or residual stress accumulation. Furthermore, as the build geometry evolves, variations in feature dimensions and heat accumulation alter the local thermal history, often necessitating dynamic adjustments to processing parameters [24]. To address these challenges, modern PBF-LB platforms integrate in-situ monitoring and adaptive control capabilities, enabling real-time detection of process anomalies and on-the-fly optimization of scan strategies. Continued progress in monitoring and closed-loop control technologies will be pivotal to unlocking PBF-LB's full potential for delivering consistent, high-integrity, and application-ready metal components [25]. To meet these demands, modern PBF-LB platforms are increasingly outfitted with in-situ detectors that continuously monitor the process zone, capturing digital information as the build progresses [26]. Depending on the sensing strategy, these systems may employ on-axis configurations, in which detectors share the optical path with the processing laser to capture signals such as optical back-reflection or photodiode-based emissions, or off-axis configurations, where external sensors including optical assemblies, cameras, or microphones observe the melt pool and powder bed from fixed vantage points [27–29]. Each approach offers distinct advantages: on-axis detectors provide precise, direct measurements of the laser–material interaction, while off-axis sensors enable a wider spatial perspective of melt pool behavior and surrounding powder dynamics [30, 31]. Increasingly, PBF-LB systems combine multiple sensing channels such as optical, acoustic, thermal, and even high-speed imaging—operating in synchrony to deliver complementary and time-resolved views of the transient, multi-physics phenomena that govern part quality [32]. This integrated, multi-sensor perspective forms the basis for advanced, real-time monitoring and adaptive control strategies [33–35].

The integration of advanced sensor monitoring systems into PBF-LB has substantially accelerated the adoption of machine learning (ML) for in-situ process control and quality assurance [36]. Modern ML approaches—particularly deep learning architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and transformer-based models—leverage high-dimensional, time-series data from optical, thermal, and acoustic sensors to capture the complex, non-linear dynamics of the PBF-LB process [37–39]. These models surpass traditional statistical techniques (e.g., principal component analysis, handcrafted feature extraction) by autonomously learning latent representations of process anomalies, including keyhole porosity, lack-of-fusion defects, and spatter-induced irregularities [39, 40]. By analyzing spatially and temporally resolved process signals such as melt pool emissions, plume spectroscopy, or layer-wise photodiode measurements—deep learning models can identify subtle, transient deviations that often precede defect formation [41, 42]. Supervised learning frameworks, when trained on synchronized in-situ monitoring data and ex-situ characterization (e.g., X-ray CT, SEM), enable real-time defect classification, while unsupervised methods (e.g., autoencoders, variational inference) detect previously unseen process anomalies without the need for labeled datasets [43–45]. Reinforcement learning (RL) extends these capabilities by dynamically adjusting processing parameters (e.g., laser power, scan speed) through closed-loop control, adapting to evolving process conditions [46, 47]. The synergy between multi-sensor fusion and advanced ML not only enhances process stability but also moves PBF-LB closer to fully autonomous systems capable of self-optimization and real-time quality certification [48]. Graph-based studies further show that PBF-LB monitoring can benefit from embedding spatial and physical process priors into the learning architecture, with physically informed Graph Neural Networks (GNNs) improving acoustic defect classification over conventional CNNs and time-correlation graph structures [49]. Recent physics-informed and physics-guided ML studies also show that manufacturing models benefit from embedding process-relevant descriptors and physical interpretation into data-driven learning, rather than relying on black-box prediction alone [50–52]. While deep learning models have demonstrated strong performance in PBF-LB process monitoring, their deployment in real-time manufacturing environments presents additional challenges. Many state-of-the-art architectures achieve high accuracy at the expense of large parameter counts and substantial computational requirements, which can hinder implementation in edge-computing or resource-constrained settings typical of industrial PBF-LB machines. High model complexity increases inference latency, demands greater memory bandwidth, and can complicate integration with existing machine controllers. Consequently, there remains a critical need for lightweight architectures that retain predictive accuracy while minimizing computational overhead. Addressing this trade-off between model performance and computational efficiency is essential for advancing practical, in-situ ML solutions in PBF-LB.

The push toward newer material compositions in PBF-LB—particularly multi-material and compositionally graded structures—further amplifies the need for robust process monitoring strategies. In high-value engineering sectors, the ability to integrate disparate material properties within a single component enables unprecedented functionality for thermal management, wear resistance, and structural optimization [53–56]. The steel–Cu multi-material system, for example, is of considerable industrial interest due to its combination of mechanical strength, oxidation-corrosion resistance (stainless steel), and high thermal-electrical conductivity (Cu and its alloys), making it attractive for applications such as heat exchangers, mold tooling, and multifunctional structural components [57–59]. Laser processing of steel-copper multi-material structures presents significant challenges due to the stark contrast in their thermophysical properties (melting point, coefficient of thermal expansion and thermal conductivity). This mismatch, together with the complex thermal cycles of PBF-LB processing, leads to the formation of residual stresses and promotes liquid metal penetration cracking, wherein molten copper penetrates austenitic steel grain boundaries and disrupt their coherency [60]. Fe-Cu phase diagram presents a miscibility gap, leading to the separation of the initially homogeneous liquid into two separate liquids (one Fe-rich and one Cu-rich) upon solidification [61]. Furthermore, the high reflectivity of copper and its alloys at the near-infrared wavelengths (700–2500 nm) typically used in conventional PBF-LB systems reduces laser absorption, exacerbating porosity and compromising build quality. Powder blending and the use of advanced feeding systems—such as multi-nozzle or multi-hopper setups and double coaters [62]—have shifted the focus of additive manufacturing research toward the development of functionally graded structures (FGSs). In these systems, the transition between dissimilar alloys is achieved through multiple intermediate compositions. This strategy aims to mitigate cracking phenomena at material interfaces, either by introducing filler alloys with high mutual miscibility (e.g., using Ni and its alloys as interlayers between steel and copper alloys [63]) or by employing carefully selected compositional gradients. Beese et al. [64] investigated the fabrication of Ti-6Al-4V/316L stainless steel graded structures using intermediate V-rich compositions, demonstrating that properly designed compositional steps can prevent the formation of brittle intermetallic phases. Beyond suppressing cracking, the use of tailored interlayers also serves as an engineering tool for microstructural control and texture manipulation. In a previous study [65], the authors examined the 316L–CuCrZr system in the form of functionally graded materials. It was shown that intermediate compositions containing 10–40 wt.% CuCrZr led to pronounced texture development along the build direction, while compositions between 40 and 90 wt.% CuCrZr promoted grain refinement and ferrite formation.

Despite recent advances, process monitoring for multi-material PBF-LB, particularly in systems combining alloys with markedly different physical properties, remains in its infancy. Variations in melting point, thermal conductivity, and optical absorptivity within a single build can significantly influence melt pool dynamics, drive compositional segregation, trigger phase incompatibilities, and introduce defect modes such as cracking or porosity that manifest unevenly across graded regions [66]. These challenges highlight the need for richer sensorization strategies that can capture a broader spectrum of process emissions with higher fidelity. In such scenarios, reliance on a single sensing modality is insufficient; integrating complementary optical, acoustic, and thermal data streams is essential to fully capture process variability and material behavior. Addressing this gap, our previous work showed that temporally aligned AE and OE signals can be fused using CNN-based

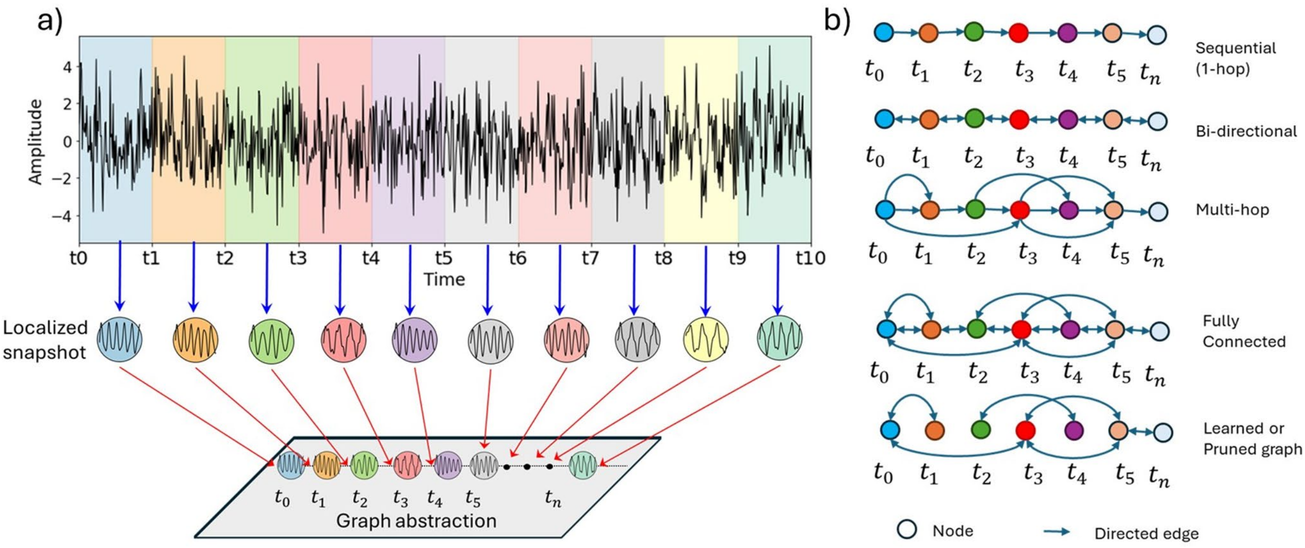

Fig. 1 **a** Time-series to Graph abstraction and **b** Edge design strategies

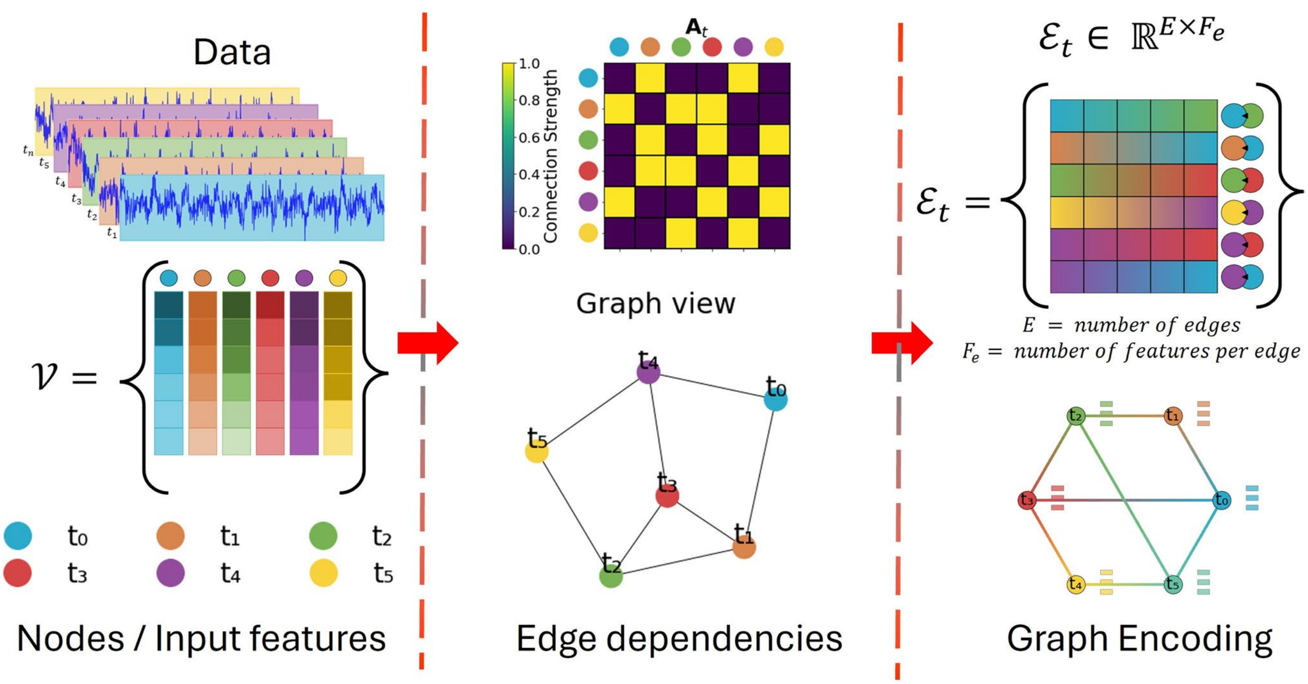

Fig. 2 Graph encoding from temporal signal features with edge attribute construction

contrastive learning to discriminate nominal 316L–CuCrZr compositional transitions in multi-material PBF-LB [67]. However, that study was primarily fusion-centric, demonstrating composition separability in a latent feature space rather than identifying how each sensing modality contributes to the decision. The present work is motivated by a different objective: establishing modality dominance and sensor ranking for composition-sensitive monitoring, which is essential for optimizing sensor configurations in practical and resource-constrained PBF-LB environments. In this context, sensor fusion is not merely beneficial but imperative, since no single modality consistently captures composition-dependent process signatures across graded regions. To move beyond fusion accuracy alone, the proposed framework replaces CNN-based contrastive embedding with learnable channel-wise shapelets and Graph Attention Network (GAT)-based temporal reasoning. This enables explicit analysis of modality saliency, shapelet activation, node-wise importance, and latent-space continuity, thereby converting multimodal fusion from a black-box classification strategy into an interpretable sensor-ranking and process-signature decoding framework Figs. 1, 2, 3, 4.

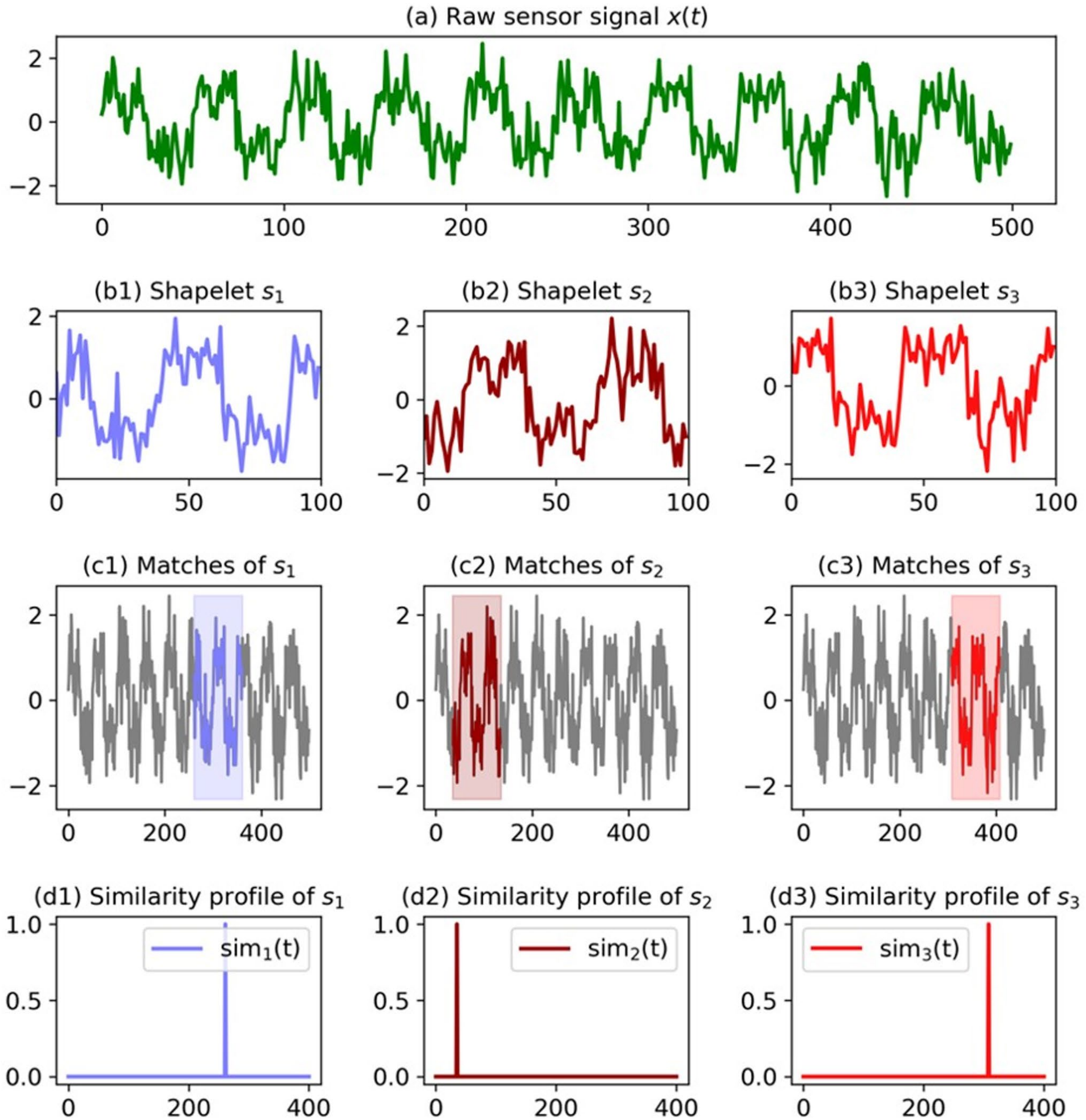

Fig. 3 Visualizing shapelet extraction: raw signal, matched segments, and similarity profiles for time-series interpretation

Central to the proposed approach is a modality-aware, lightweight ML framework that fuses AE and back-reflected OE at the laser wavelength to classify composition-dependent process signatures in multi-material PBF-LB (316L–CuCrZr), combining learnable shapelets (explainable, temporally localized features) with GAT to capture both local and long-range temporal dependencies. This design enables the model to identify modality-specific signatures of compositional variation while maintaining a compact architecture of ~4,000 trainable parameters. This compact parameter footprint motivates future evaluation for resource-constrained monitoring scenarios. By combining physical interpretability, multi-sensor fusion, and architectural efficiency, the proposed Shapelet–GAT framework offers a scalable pathway toward physics-informed, in-situ monitoring and adaptive control in complex, multi-material PBF-LB systems. The remainder of this manuscript is structured as follows: Sect. 1 provides an overview of the

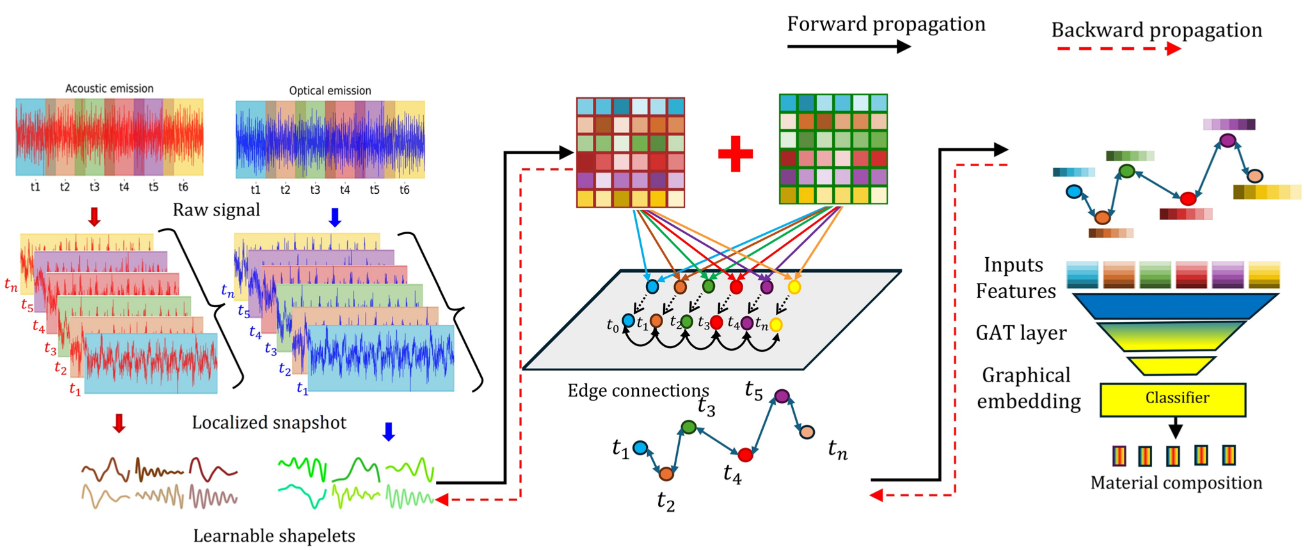

Fig. 4 Overview of the shapelet-GAT framework for multimodal time-series classification

multi-material PBF-LB process, emphasizing its relevance for fabricating compositionally graded alloys and the necessity of high-fidelity, real-time process monitoring. Section 2 details the proposed methodology, including the formulation of the graph-based learning framework, the design of the learnable shapelet extraction module, and the integration of attention-driven classification mechanisms. Section 3 describes the experimental platform, sensor configuration, and synchronized data acquisition pipeline. Section 4 presents the exploratory data analysis, incorporating the preliminary statistical characterization of the dual-modality signals and identifying composition-dependent temporal and spectral signatures in acoustic and optical emissions. Section 5 introduces the Shapelet– GAT architecture, encompassing input representation, channel-wise shapelet encoding, and attention-based message-passing mechanisms for temporal graph learning. Section 6 reports and interprets the multi-material composition-monitoring results, supported by quantitative performance metrics and interpretability analyses. Finally, Sect. 7 concludes the study by summarizing the principal findings and outlining future research directions.

2 Methodology

2.1 Graph neural network

Graph Neural Networks offer a powerful and principled framework for modeling complex dependencies in data represented as graphs. They are particularly effective in capturing structured relationships in time-series data, especially in situations where traditional sequential models fall short in representing long-range dependencies or multi-context interactions or sensing modalities. Rather than treating time-series as flat sequences, GNNs reinterpret the data as a graph in which each node corresponds to a meaningful segment of the signal such as a fixed-size temporal window, a feature embedding, or a latent representation extracted from an encoder. Edges connect nodes that share relevant relationships, which may be defined based on temporal proximity, similarity in feature space, co-occurrence patterns, or learned dependencies. These edges may reflect direct signal transitions, correlation strength, or domain knowledge. This graph-based view allows the model to reason about interactions not only between adjacent segments but also across non-local or contextually linked regions of the signal.

Within a GNN, each node's representation is updated by aggregating information from its neighbors, encoding both structural and feature-level information. This mechanism makes GNNs highly effective for tasks requiring relational context, such as time-series fusion, structural pattern analysis, or dynamic behavior modeling. In GNNs, each node's representation is iteratively updated by aggregating features from its neighbors. This process encodes both local structure and node-specific attributes, making GNNs particularly suitable for tasks involving relational or spatial context—such as time-series fusion, temporal waveform analysis, or sensor-based monitoring. Formally, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph with node set $\mathcal{V}$, edge set $\mathcal{E}$, and initial node features $h_v^{(0)} \in \mathbb{R}^d$ for each node $v \in \mathcal{V}$. GNNs operate by iteratively updating node representations through a message passing mechanism: each node aggregates information from its neighbors, incorporating both its own features and those of structurally relevant peers. This process is mathematically described as:

$$h_v^{(l+1)} = \sigma\left(W^{(l)} \cdot AGG^{(l)}\left(\left\{h_u^{(l)} | \mu \in \mathcal{N}(v) \cup \{v\}\right\}\right)\right) \tag{1}$$

Here, $h_v^{(l)}$ denotes the feature embedding of node v at layer l, $\mathcal{N}(v)$ is the set of neighboring nodes, $AGG^{(l)}\bullet$ is a permutation-invariant aggregation function such as sum, mean, or attention, $W^{(l)}$ is a learnable weight matrix, and $\sigma(\bullet)$ is a non-linear activation function.

When attention mechanisms are employed such as in GATs, the aggregation is weighted by the learned importance of each neighbor, allowing the model to emphasize the most informative connections. This makes GNNs particularly well-suited for time-series applications where patterns may span variable lengths, exhibit complex dependencies, or occur across irregular temporal intervals. The recursive aggregation mechanism ensures that each node's embedding captures higher-order structural information across multiple layers. In tasks involving multimodal signals or segmented temporal data, each node may represent a localized signal snapshot, while the edges capture similarity, correlation, or transition dynamics.

2.2 Shapelets for time-series

In sensor-based time-series analysis encompassing applications such as structural vibration monitoring, AE diagnostics, photonic signal interrogation, and physiological biosignal processing, the recorded signals often contain class-discriminative patterns that are temporally localized and embedded within noise-prone, non-stationary sequences. Conventional methods relying on global statistical descriptors, spectral decompositions (e.g., Fourier or wavelet transforms), or time-averaged features may fail to capture subtle, transient motifs that are critical for identifying specific operational states, fault signatures, or physiological anomalies. To address these limitations, shapelets have been proposed as short, discriminative subsequences that optimally differentiate between classes. By focusing on localized temporal structures, shapelet-based methods offer a principled and interpretable framework for fine-grained classification and pattern discovery in sensor-rich environments.

Formally, a shapelet $s \in \mathbb{R}^L$ is a local template of length L that captures recurring and class-relevant patterns within a longer time-series $x \in \mathbb{R}^T$, where typically $T \gg L$. These motifs may represent localized phenomena such as transient sensor activations, microstructural deformation responses, or modality-specific signal bursts, which may vary in duration, phase, or amplitude across different mechanisms. To quantify the similarity between a shapelet s and a sensor signal x, we compute the minimum $z-$ normalized squared Euclidean distance between the shapelet and all contiguous subsequences of the signal.

$$D(x,s) = \min_{\mathrm{t}\,\in\{1,\ldots,T-L+1\}} \frac{1}{L}\left\|\widehat{x}_{t:t+L-1} - s\right\|^2 \tag{2}$$

where $\widehat{x}_{t:t+L-1}$ denotes the $z-$ normalized version of the windowed segment of the signal starting at index t, and s is the unnormalized shapelet. This formulation ensures robustness to variations in signal baseline and amplitude while preserving the original form of the learned shapelet, thereby supporting both discriminative performance and interpretability. Recent advances have extended this formulation by making shapelets learnable parameters within an end-to-end optimization framework. In these models, shapelets are initialized randomly or heuristically, and are subsequently optimized using gradient-based learning to minimize a task-specific objective, such as cross-entropy for classification.

3 Methodology

The graph-based learning methodology is proposed for classifying synchronized dual-channel time-series originating from two different sensing modalities: AE and OE's in the range of laser wavelength. The method involves three main stages: temporal graph construction, shapelet-based feature extraction, and graph-based classification. Initially, each time-series signal is normalized to zero mean and unit variance to ensure numerical stability. The standardized signal is segmented into overlapping windows using a fixed-size sliding window; each segment is treated as a graph node. Temporal continuity is maintained by connecting adjacent windows with bidirectional edges, forming a graph that captures local temporal dependencies. In the second stage, the model learns channel-specific shapelets that encode discriminative local patterns. A fixed set of shapelets is optimized per modality, capturing characteristic waveform features relevant to the task. For each node, the shapelet module computes the minimum distance between every shapelet and all possible subsequences within the window. These distances are concatenated to form a compact, interpretable node representation that replaces the original input, enriching the graph with modality-aware features. In the final stage, shapelet-encoded graphs are processed by a stack of GAT layers to model temporal dependencies and enrich node-level representations. The initial layers employ multi-head attention to enhance training stability and capture diverse contextual interactions among neighboring nodes. Attention scores are computed across edges, allowing the model to adaptively weigh contributions from adjacent time segments based on their contextual relevance. Each GAT layer is followed by batch normalization, non-linear activation, and dropout to improve generalization and mitigate overfitting Fig. 5.

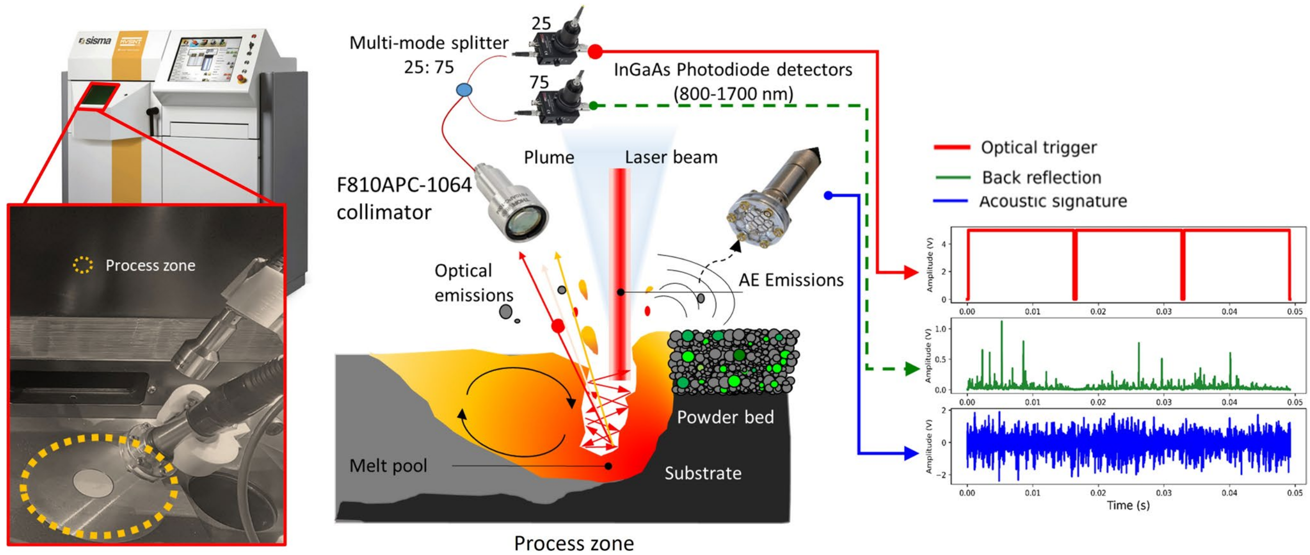

Fig. 5 Schematics of the proposed sensing system comprising a photodiode and AE sensors

Following attention-based message passing, the refined node embeddings are aggregated using mean pooling to obtain a graph-level representation that captures both localized waveform characteristics and higher-order temporal structure. This global embedding is then passed through a downstream classification head to generate the final prediction. The architecture supports unified training, with shapelet parameters and GAT weights optimized cohesively through a single training loop. Additionally, the shapelet module enables per-channel attribution, facilitating interpretation of modality-specific contributions (e.g., AE vs. OE) through selective masking or distance-based relevance analysis.

Figure 4 illustrates the Shapelet-GAT pipeline. Standardized dual-channel signals are segmented into overlapping windows, each forming a node in a temporal graph with bidirectional edges. Shapelet distances are computed per node to generate compact features that replace raw inputs. The resulting graph is processed by stacked GAT layers for attention-based information propagation. A global embedding is obtained via mean pooling and passed to a classifier for final prediction. The entire model is trained end-to-end using cross-entropy loss and Adam optimizer, with validation accuracy guiding selection.

3.1 Experimental setup, materials, data acquisition and methodology

3.1.1 Experimental setup and materials

To enable high-resolution in-situ diagnostics of the PBF-LB process, a synchronized multi-sensor monitoring framework was implemented, combining OE and AE sensing. This multi-modal system was designed to temporally correlate photonic and phononic signatures of dynamic process events with microsecond-level precision. Optical emissions from the laser–material interaction zone were collected using a fixed-focus F810APC-1064 collimator (Thorlabs), aligned for direct melt pool observation. The optics employ a narrowband AR coating centered at the 1064 nm band, providing high transmission and sensitivity to back-reflected laser radiation from the process zone. The optical signal was routed through a 25:75 multimode fiber coupler, where 25% of the signal was directed to a PDA20CS2 photodiode (800–1700 nm) and amplified to a saturation level of + 5 V, enabling it to serve as a robust trigger signal. The remaining 75% of the signal was directed to a second PDA20CS2 for continuous intensity quantification, allowing dynamic tracking of process-zone radiative emissions. Owing to the fixed-focus F810APC-1064 collimator, the photodiode readings are primarily sensitive to the wavelength band transmitted through its 1064 nm-optimized optics, enhancing responsiveness to OE near the laser wavelength. Simultaneously, process-induced acoustic activity was captured using a CM16/CMPA broadband airborne AE sensor (Avisoft Bioacoustics), with a flat frequency response spanning 2–150 kHz. The sensor was positioned to maximize sensitivity to airborne acoustic transients associated with spatter events, melt pool instabilities, and subsurface defect formation. All sensor signals were digitized using an Advantech PCIe-1840 data acquisition card, operating at a 400 kHz sampling rate and ± 5 V dynamic input range, ensuring compliance with the Nyquist criterion for full AE bandwidth. A threshold-based triggering mechanism initiated synchronized acquisition when the optical trigger signal exceeded 0.5 V, producing characteristic square-wave signals (> 5 V)

Table 1 Chemical composition of the 316L powder feedstock provided by Oerlikon

Weight percent (nominal)					
Fe	Cr	Ni	Mo	C	Other
Balance	18	12	3	< 0.03	< 1.0

Table 2 Chemical composition of the CuCrZr powder feedstock provided by Eckart

Weight percent (nominal)				
Cu	Cr	Zr	Fe	Si
Balance	0.96	0.12	0.021	< 0.004

that precisely demarcated the temporal window of each scan track. The system was deployed on a SISMA MySint 100 PBF-LB machine, equipped with a 1070 nm fiber laser, 55 µm laser beam focal point diameter, and maximum output power of 200 W, operating in an argon-purged chamber at atmospheric pressure. Additional implementation details of the experimental rig (e.g., mounting geometry, optical routing, and synchronization circuitry) are reported in our prior work [67]; here we summarize only the elements necessary for the present study.

This study systematically investigated the fabrication of functionally graded steel–copper structures via PBF-LB using gas-atomized 316L stainless steel powder provided by Oerlikon Metco (composition given in Table 1) and Cu-1Cr-0.3Zr (wt.-%), hereafter referred to as CuCrZr, alloy powder provided by Eckart GmbH (composition given in Table 2). Overall 6 distinct compositions have been used: pure 316L stainless steel, pure CuCrZr and 4 316L-CuCrZr mixtures (80:20, 60:40, 40:60, 20:80 316L:CuCrZr wt.%). For premixing the powders, a Turbula dry powder blender has been used, operated for 1h at 46 rpm, in order to homogenise 200 gr of each premixture separately.

The 316L-CuCrZr multi-material sample was built in the form of a functionally graded structure. The first 50 layers of the structure were composed of pure 316L stainless steel powder. The next 50 layers (51–100) were built using the 20 wt.% CuCrZr powder, succeeded by the 40 wt.% CuCrZr powder (101–150 layers).The process was repeated with gradually increasing the CuCrZr content by 20 wt.% for every new 50 layers, until pure CuCrZr was reached, resulting in a final structure with a total size of 1 × 1x0.85 cm3. Two laser parameter sets were utilized: a 316L-optimized condition with a laser power of 125 W, scan speed of 750 mm/s, and layer thickness of 20 µm was used for the pure stainless steel builds; meanwhile, a unified set of parameters with 200 W power, 600 mm/s scan speed, and a 30 µm layer thickness was applied to all Cu-containing premixtures. The selected parameter set yielded a constant volumetric energy density (VED) of 123 J/mm3. This VED was fixed for all steel–copper compositions to isolate the influence of alloy composition, eliminating confounding effects from variations in processing parameters. Maintaining a uniform energy input ensured that differences in melt pool dynamics, defect formation, and emission signatures could be directly attributed to compositional changes rather than thermal input. This consistency was essential for establishing a reliable correlation between in-situ sensor responses and material composition, forming the basis for the multimodal learning framework proposed in this study. The process parameters have been summarized in Table 3.

Inert argon atmosphere (100 ppm oxygen content) is used upon printing to minimize oxidation. To prevent cross-contamination, the PBF-LB system was thoroughly cleaned following the fabrication of each composition, and the appropriate premixture was freshly loaded into the powder supply unit. Since SiSMA MySINT 100 PBF-LB is equipped with a single powder feeder, the powder replacement has been done manually. After printing, the specimen was removed from the build plate and cut into two vertical cross sections using Electrical Discharge Machining (EDM) to preserve structural integrity and avoid introducing mechanical artifacts. Post-process specimen removal was followed by standard metallographic preparation, including grinding with SiC abrasive papers from 600 up to 4000 grit. Subsequently, for polishing diamond suspensions of 3 µm and 1 µm have been used, followed by final polishing through vibratory polishing with a 0.05 µm alumina suspension for 10 h. In order to reveal the melt-pool formation, cross sections were etched, using 1.5% Nital etch agent. For microscopy analysis, a TESCAN MIRA 3 field emission gun scanning electron microscope (FEG SEM) was employed.

Table 3 PBF-LB process parameters on building the cubic structure

	Laser power (W)	Scan speed (mm/s)	Hatch distance (µm)	Layer thickness (µm)	Label
316L wt.%	125	750	80	20	100%-316L
80–20 316L-CuCrZr wt.%	200	600	90	30	20%-Cu
60–40 316L-CuCrZr wt.%	200	600	90	30	40%-Cu
40–60 316L-CuCrZr wt.%	200	600	90	30	60%-Cu
20–80 316L-CuCrZr wt.%	200	600	90	30	80%-Cu
CuCrZr wt.%	200	600	90	30	100%-Cu

3.1.2 Sample characterization

Micrographs of the samples with different compositional gradients are presented in Fig. 6. For the 316L part, dense microstructures without cracks and only a few pores are observed. This is expected, as 316L is known for its excellent processability in PBF-LB. In contrast, pronounced cracking is observed in the 20 wt.% and 40 wt.% CuCrZr mixtures, whereas the 60 wt.% and 80 wt.% CuCrZr mixtures remain crack-free. Pure CuCrZr layers are again free of solidification cracks; however, they show a significant number of lack-of-fusion defects. Overall, the amount of porosity increases with higher CuCrZr content. The observed evolution of cracking and porosity with CuCrZr fraction follows trends reported in our previous publication [65]; the present work leverages these known material states to investigate their signatures in in-situ monitoring data.

Increasing CuCrZr content leads to a progressive reduction in melt pool depth. In Fig. 7, the melt pool boundaries for 20, 40, 60, and 80 wt.% CuCrZr are marked with black arrows. At 20 and 40 wt.% CuCrZr, the melt pools are relatively deep, with depths of ~ 108 µm and ~ 64 µm, respectively, and the width-to-depth ratio indicates processing in the keyhole or transition mode. With higher CuCrZr contents, a transition to conduction mode is observed, as the melt pool depth decreases to ~ 40 µm at 60 wt.% and ~ 34 µm at 80 wt.% CuCrZr. This trend is attributed to the higher reflectivity and thermal conductivity of Cu, which reduce melt pool stability and promote rapid solidification. In the 20 and 40 wt.% CuCrZr samples, large cracks aligned parallel to the build direction along the melt pool center, as well as smaller, irregular cracks, can be observed. These compositions also show a pronounced <100> texture along the build direction, whereas the 60 and 80 wt.% mixtures exhibit an almost random texture. As reported previously [65], the reduction in melt pool depth correlates with the weakening of texture intensity.

The large cracks observed are likely solidification cracks, caused by CuCrZr forming liquid films ahead of the already solidified 316L. The occurrence of smaller cracks, and their disappearance at higher CuCrZr fractions, is consistent with previous studies on 316L–CuCrZr functionally graded structures, which identified ~ 50 wt.% CuCrZr as the threshold beyond which micro-cracking is suppressed. This transition to crack-free microstructures is attributed to a change in the solidification sequence in CuCrZr-rich mixtures (50–90 wt.% CuCrZr), driven by the miscibility gap that produces two immiscible liquid phases: an Fe–Ni–Cr-rich liquid and

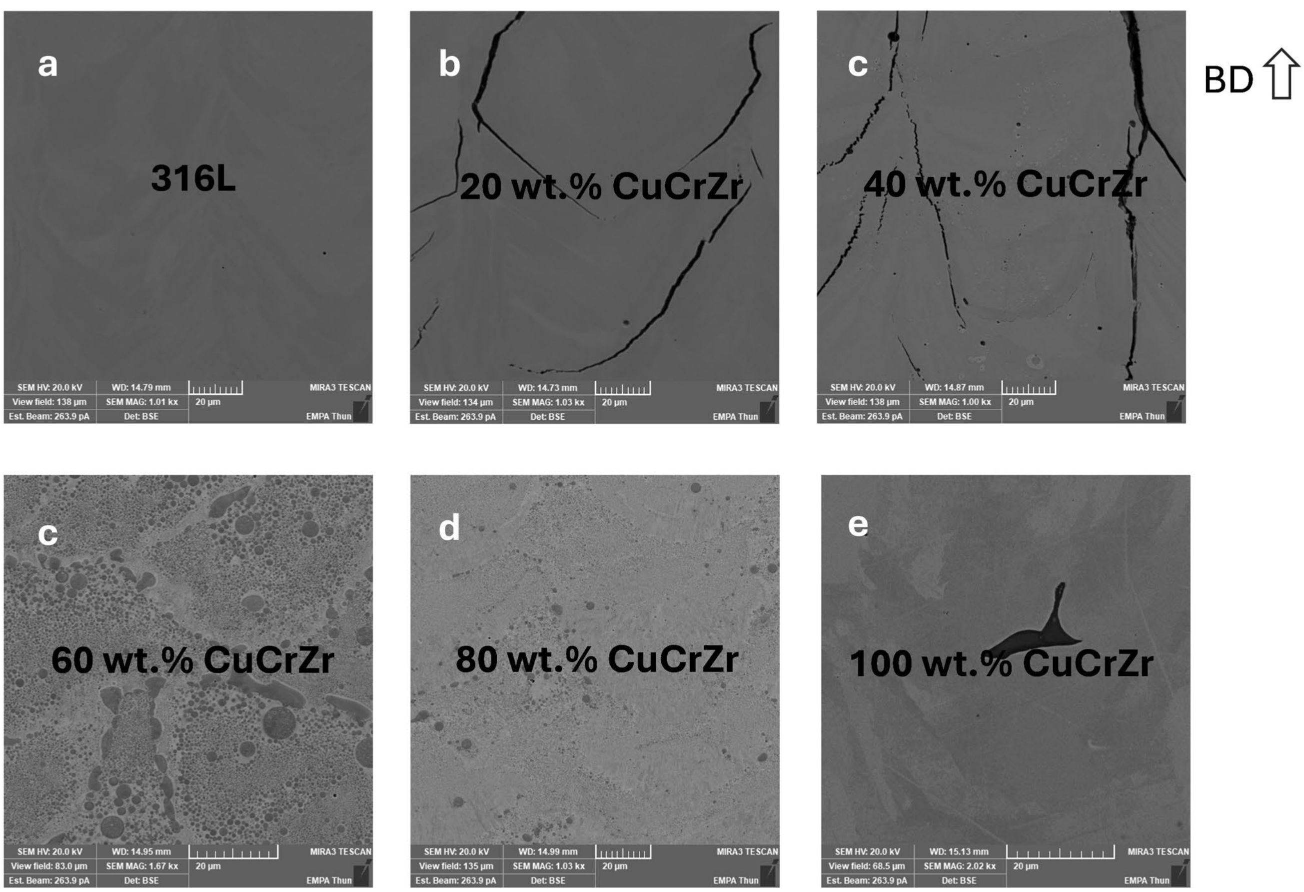

Fig. 6 Micrographs of the different compositions of the structure

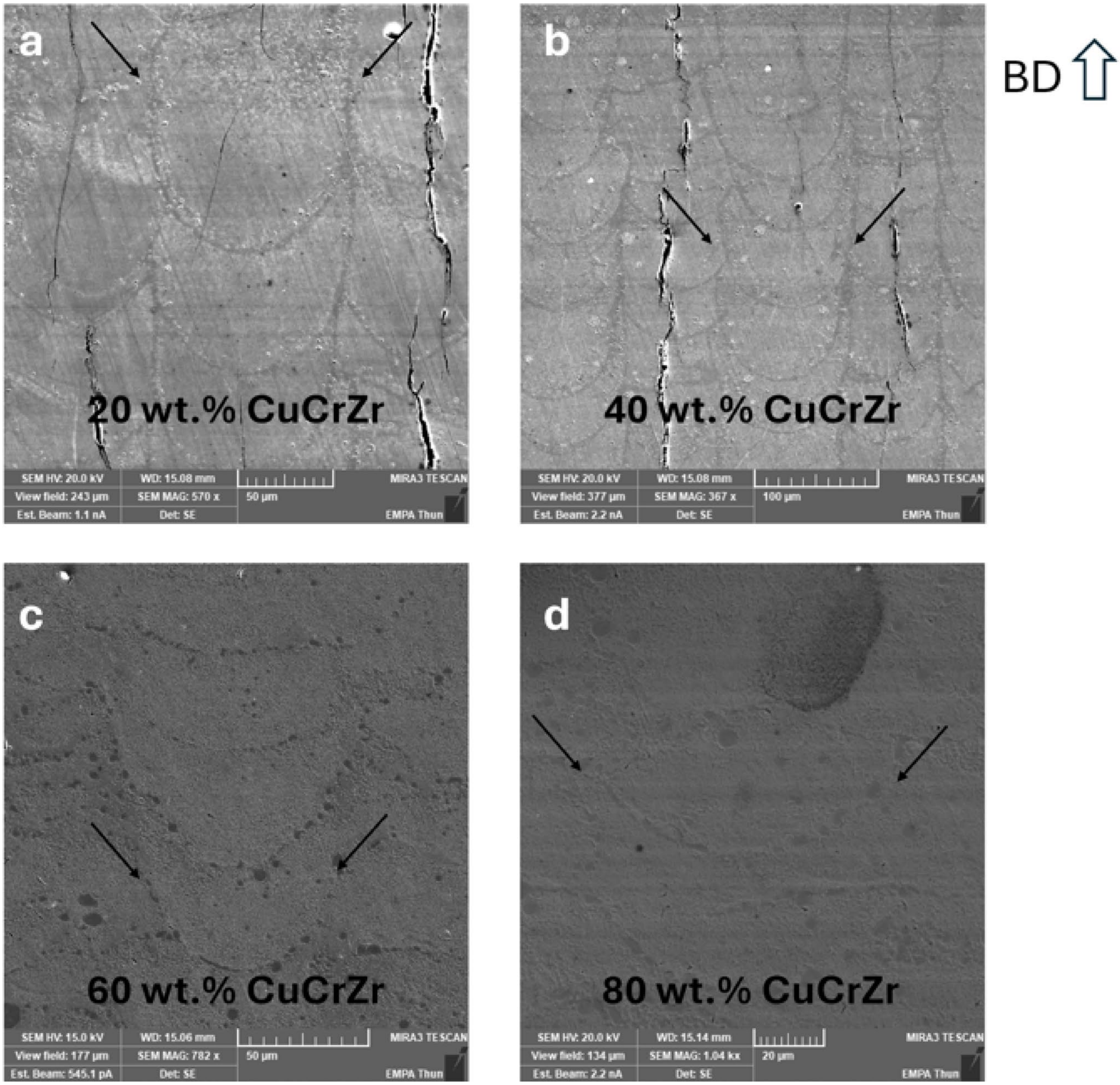

Fig. 7 Melt pool morphologies for different mixtures across the building direction

a Cu-rich liquid. The phase separation is accompanied by diffusion of 316L alloying elements—particularly Ni—which alters the primary solidification path, leading to the formation of BCC ferrite alongside FCC austenite [65]. The presence of α-ferrite improves grain boundary cohesion and reduces Cu wettability compared to γ-austenite, thereby suppressing intergranular Cu infiltration and mitigating crack formation.

Thus, despite changes in thermal input and melt pool geometry with increasing CuCrZr content, the observed transition from crack-susceptible to crack-resistant behavior above ~ 60 wt.% CuCrZr is assumed to be composition-driven. A more detailed analysis of the underlying cracking mechanisms, however, lies beyond the scope of this study.

In addition to the observed transition from crack-susceptible to crack-resistant behavior with increasing CuCrZr content, variations in melt pool depth and morphology across the compositional gradient are also expected to influence the process monitoring signals. Shallower melt pools in CuCrZr-rich regions alter melt pool dynamics and solidification rates, which in turn affect the frequency spectrum and amplitude of process-generated acoustic waves. At the same time, the higher reflectivity and modified absorptivity of Cu-rich compositions change the spatial and temporal distribution of OE's from the process zone. Captured simultaneously, these multimodal process signatures provide indirect yet highly sensitive indicators of the underlying composition, enabling a reliable distinction between CuCrZr-rich and 316L-rich regions. This motivates the use of multimodal AE and OE signals for composition-sensitive monitoring in multi-material PBF-LB. In this study, the analysis is limited to distinguishing the planned nominal composition regions. Detecting unintended composition drift or off-nominal deviations requires future validation with deliberately introduced composition changes.

3.2 Data acquisition pipeline and dataset

During the layer-by-layer fabrication of the compositionally graded cube, OE's from the laser–powder interaction zone consistently exceeded the 0.5 V threshold on Channel 1 (PDA20CS2 photodiode) at the initiation of each layer, thereby triggering synchronous data acquisition across both OE (Channel 2) and AE (Channel 3, CM16/CMPA sensor) channels. The photodiode gain was dynamically adjusted in real time to maintain a saturation voltage of 5 V ($\pm$0.1 V) throughout the scanning process, ensuring high-fidelity temporal synchronization between the plasma emission intensity (captured from the 75% split of the 1070 nm optical signal) and stress wave activity (recorded in the 0–150 kHz range). The acquired signals were segmented into 12.5 ms time windows, corresponding to 5000 data points per window based on the 400 kHz sampling rate. Each segment was mapped to a distinct scan track, enabling localized, time-resolved analysis. Acoustic emission signals were post-processed using a fourth-order Butterworth low-pass filter with a cutoff frequency of 150 kHz, selected to align with the sensor's frequency response and suppress out-of-band noise. This filtering protocol was applied uniformly across all five powder compositions to ensure consistency and comparability. This synchronized multi-sensor framework yielded temporally correlated optical–AE datasets (summarized in Table 4), enabling direct, layer-specific comparisons of key process signatures. These included plasma emission fluctuations (OE) and melt pool instability frequencies (AE), captured at identical temporal resolutions. The resulting data structure provides a robust platform for investigating composition-dependent process anomalies in multi-material PBF-LB, facilitating quantitative comparisons of thermal and mechanical transients associated with each powder blend.

Table 4 The total number of 12.5 ms (5000 data points) AE and optical windows that were extracted during the fabrication of the cube [67]

Category	Windows per AE and OE channel
20%-Cu	5448
40%-Cu	6205
60%-Cu	5457
80%-Cu	5457
100%-Cu	5457

The class labels used in this study correspond to the nominal powder composition assigned to each graded region of the build. Therefore, the classification task evaluates whether the synchronized AE and OE windows carry discriminative signatures associated with these known composition classes. The present dataset does not contain independent chemical measurements for each individual scan track, nor does it include intentionally generated off-nominal composition deviations.

3.2.1 Exploratory data analysis

To investigate temporal trends and emission characteristics, representative raw time-series signals were plotted for both acoustic and optical modalities across five copper compositions: 20%, 40%, 60%, 80%, and 100% (Fig. 8). The acoustic signals predominantly exhibited oscillatory, noise-like behavior with fluctuating amplitudes, reflecting the complex, broadband nature of stress-wave emissions during laser-material interaction. In contrast, the OE signals displayed distinct transients and bursts, particularly pronounced in higher copper compositions. Notably, in the 100%-Cu condition, the OE signal exhibited sharp, high-amplitude peaks, indicating abrupt changes in reflectivity or energy coupling at the interaction zone.

3.3 AE signal processing

The frequency characteristics of the AE signals were further examined by computing the Fast Fourier Transform (FFT) via Welch's method to estimate the power spectral density (PSD) (Fig. 9). This frequency-domain analysis revealed a progressive shift in energy distribution as a function of copper content. As illustrated in Fig. 9, the dominant spectral content for all compositions is concentrated below 100 kHz, underscoring this range as the principal bandwidth for capturing AE phenomena during processing. Lower Cu compositions (20%–40%) exhibited a broader spectral spread, with significant energy extending up to 100 kHz at higher frequencies. In contrast, higher Cu compositions (80%–100%) displayed a marked concentration of energy in the lower frequency bands ($<$60 kHz), accompanied by a pronounced attenuation in the higher frequency range. Notably, the 100%-Cu samples exhibited a substantial reduction in peak spectral energy compared to other compositions. This attenuation is plausibly attributed to copper's unique optical and thermophysical properties, which influence melt pool dynamics and, in turn, the generation of AE's. These

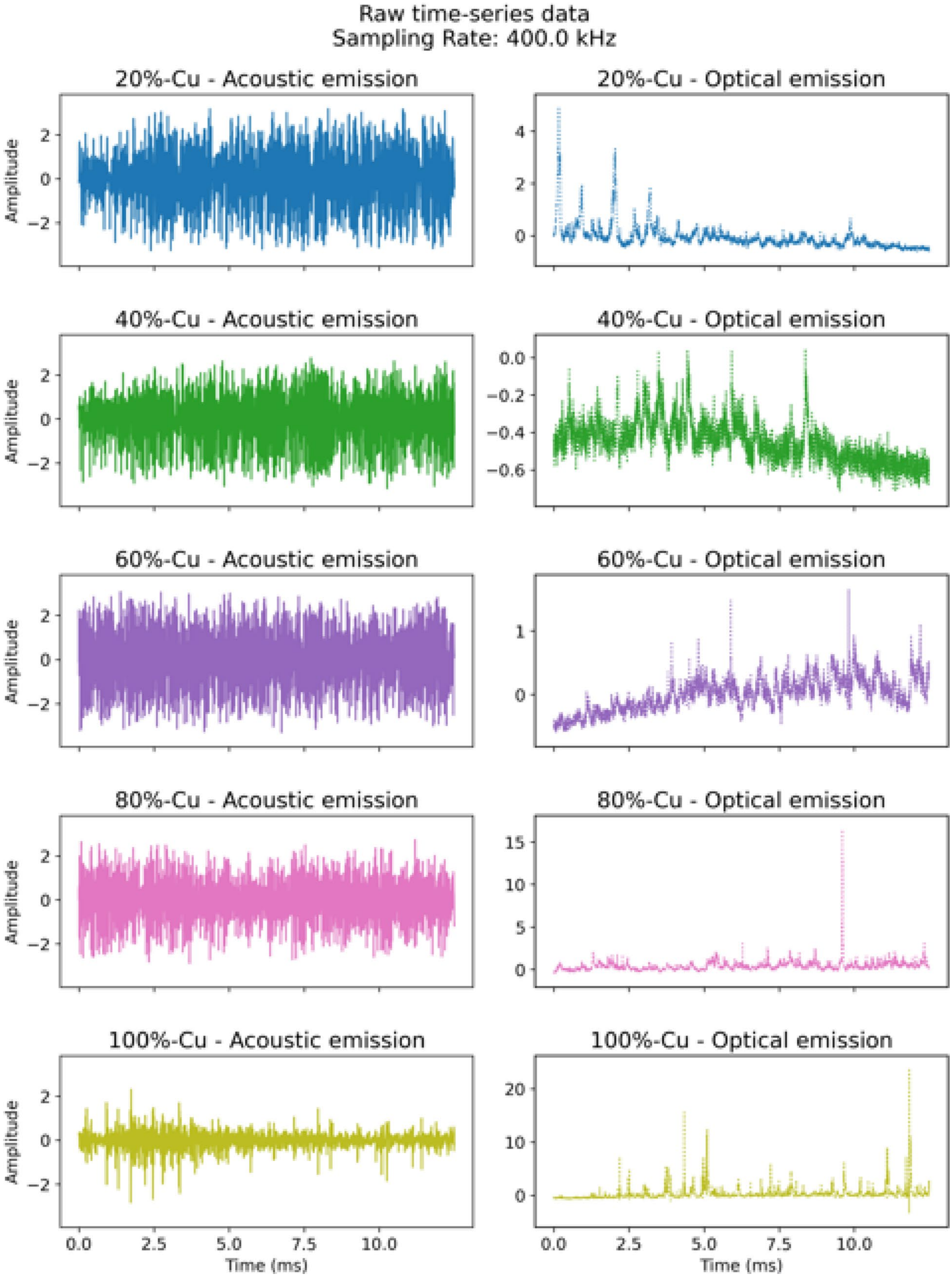

Fig. 8 Temporal profiles of AE (left) and back-reflected OE (right) for five Cu compositions (20%–100%)

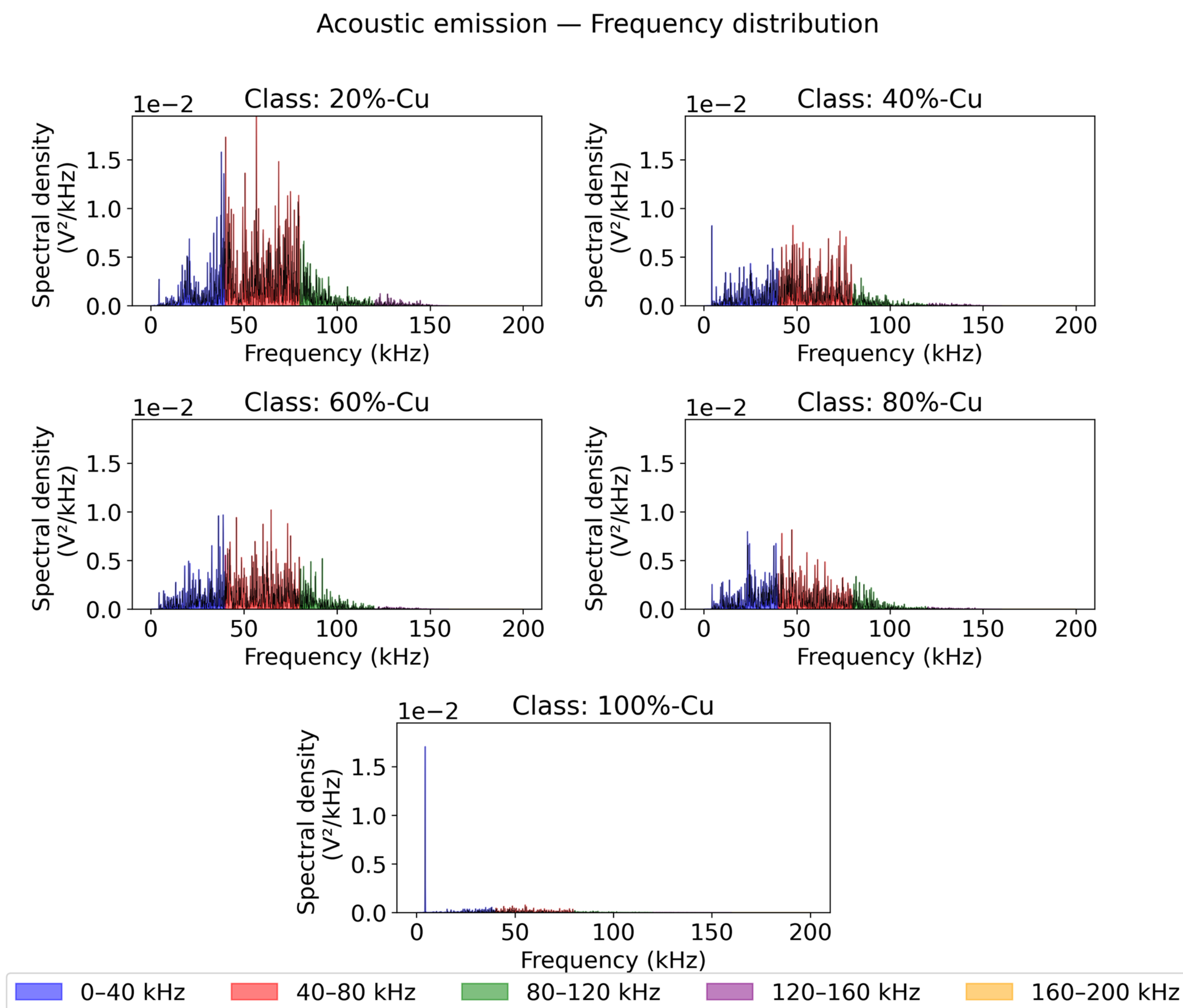

Fig. 9 Composition-dependent spectral signatures in AE signals across varying copper compositions in 316L-CuCrZr alloys

composition-dependent spectral variations likely stem from differences in energy dissipation mechanisms, acoustic wave propagation, and melt–solid coupling, all of which are inherently governed by the copper content in the alloy system.

To establish a preliminary understanding of how variations in powder composition affect the optical intensity, particularly in the spectral bands associated with the laser wavelength and its interaction zone, a frequency-domain analysis was performed on the OE. Each 12.5-ms time window was transformed to the frequency domain, and the normalised spectral energy was computed across discrete 20-kHz bands spanning the 0–180 kHz range (Fig. 10). This enabled a localised assessment of the spectral response for each compositional regime. The resulting spectral energy profiles revealed composition-dependent frequency localisation patterns. For example, the 100%-Cu samples exhibited dominant energy in the 20–40 kHz and 40–60 kHz bands, whereas mid-range compositions such as 60%-Cu demonstrated a broader and more balanced energy distribution across multiple bands. Additionally, Fig. 10 highlights regions of substantial spectral overlap among adjacent compositions, while also revealing a gradual increase in the mean normalised energy with increasing copper content. This upward trend is consistent with the high reflectivity and thermal conductivity of pure copper, which amplifies the optical return signal under laser excitation. These findings not only reinforce the influence of material properties on OE behaviour but also support the feasibility of using frequency-domain features for compositional classification and process state inference.

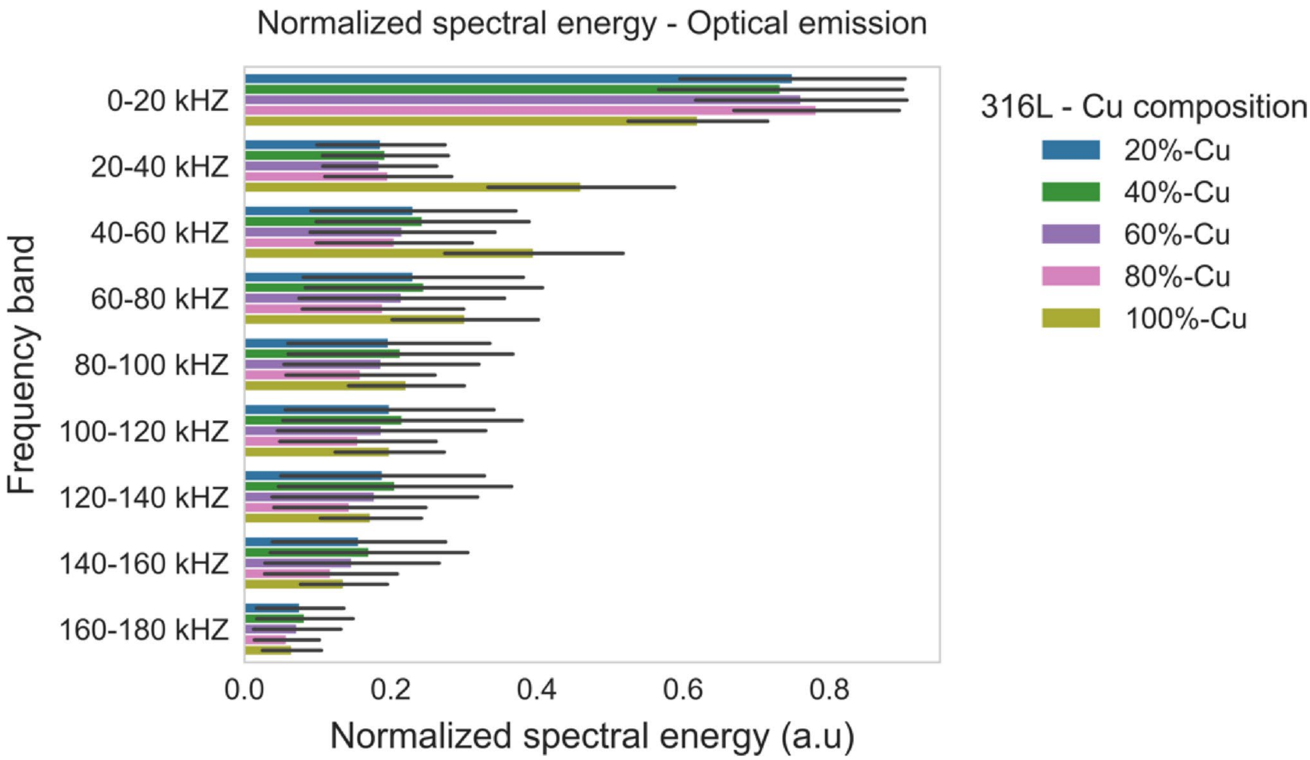

Fig. 10 Bar chart showing normalized spectral energy of back-reflected OE across nine frequency bands (0–180 kHz), grouped by Cu composition

Overall, the exploratory analysis highlights distinct temporal and spectral signatures associated with compositional variations in the 316L– CuCrZr system. These insights not only underscore the relevance of frequency-domain features in capturing material-dependent emission characteristics but also motivate the development of data-driven models capable of learning localised, modality-specific patterns for robust process monitoring and compositional classification.

4 Graph architecture

4.1 Input representation

This work proposes a graph-based learning pipeline that integrates learnable shapelet extraction with a GAT to classify multivariate time-series data originating from two synchronized sensing modalities: AE and OE (back reflection at laser wavelength), denoted as channels $\mathrm{c} \in \{1,2\}$. The pipeline begins by converting each raw time series into a sliding window graph, where the temporal structure of the signal is captured through overlapping segments. Prior to segmentation, each raw signal $\mathrm{x_c}(\mathrm{t})$ is standardized to have zero mean and unit variance, ensuring numerical stability and facilitating the joint optimization of both shapelet and attention parameters. The two-channel time series of length N is divided into overlapping windows of fixed size $W = 500$ with stride $S = 250$ resulting in M segments. Each becomes a graph node with an associated tensor $\mathrm{x_j} \in \mathbb{R}^{2\times W}$, retaining the waveform identity of both sensing channels. Next, the model learns channel-specific shapelets to extract discriminative local patterns from these windowed segments, effectively transforming raw signal windows into a more expressive latent space. Unlike traditional temporal graphs that only connect sequential windows, this work models each signal instance as a fully connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each node corresponds to a specific time window and the edge set $\mathcal{E}$ includes all possible directed edges between node pairs:

$$\mathcal{E} = \{(i,j)\,|i \neq j, i,j = 1,\ldots,M\} \tag{3}$$

This fully connected structure allows global message passing, where each node attends to all other temporal segments. Such connectivity supports the modeling of long-range temporal dependencies and global contextual patterns across the multivariate signal. The final stage involves aggregating node-level features using stacked GAT layers to compute a graph-level embedding for classification. This enables the model to leverage both local waveform features and global temporal context across both sensor channels.

Figure 11 illustrates the structure of the fully connected temporal graph. Each node represents a specific time window with features derived from both acoustic and optical

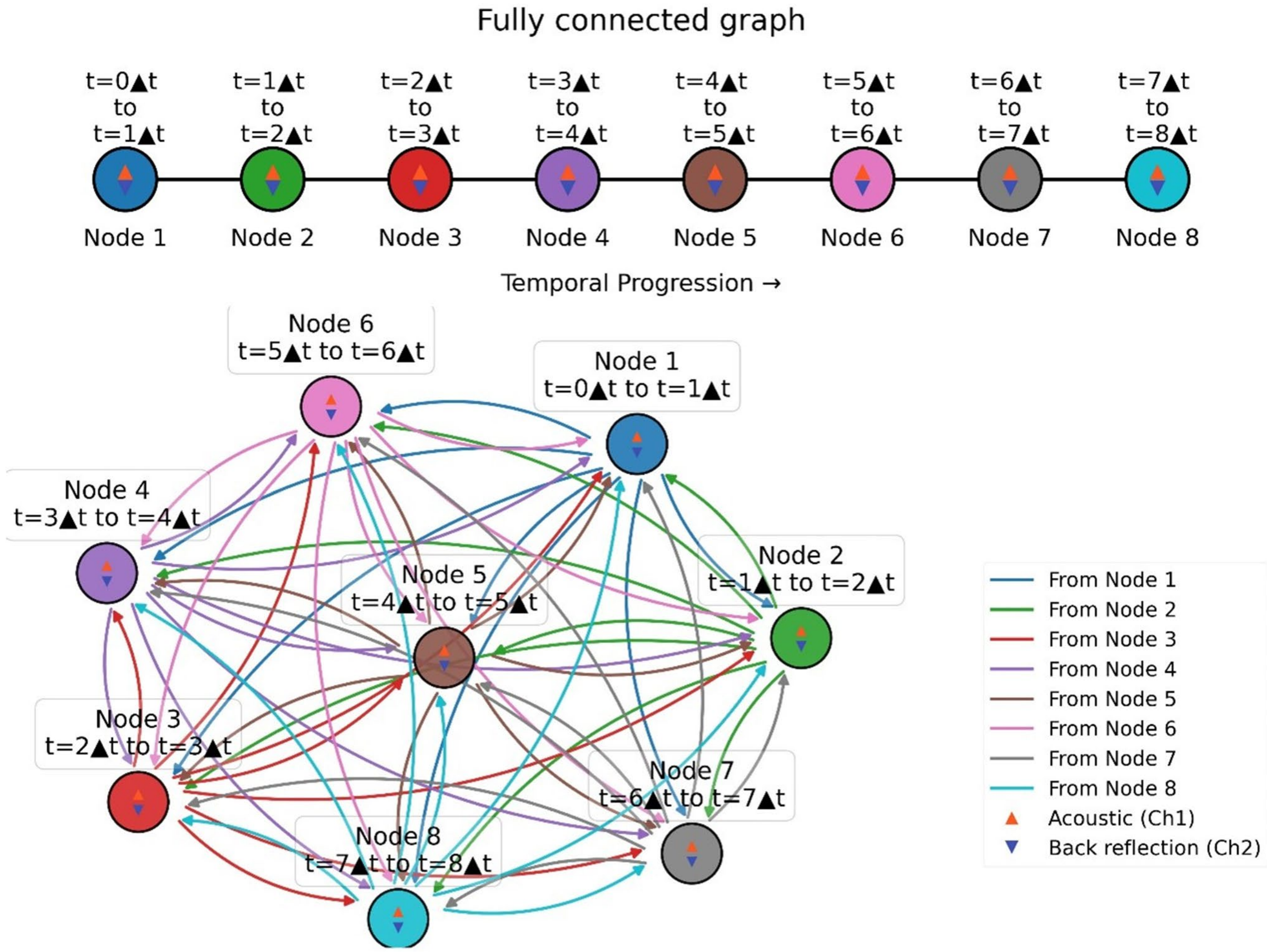

Fig. 11 Fully connected sliding window graph with multi-modal node features on eight nodes as sample illustration

channels. Directed edges between all node pairs enable comprehensive information flow across time, supporting rich cross-temporal interactions essential for effective attention-based modeling.

4.2 Learnable, channel-wise shapelet extraction

To extract discriminative temporal features from each sensing modality, we employ a learnable, channel-wise shapelet extraction mechanism applied individually to each node in the temporal graph. A total of 20 shapelets are jointly optimized during training, with 10 shapelets allocated per modality, AE and OE. Each shapelet is a short, fixed-length vector $s_k^{(c)} \in \mathbb{R}^L$, where $L = 50$ defines the temporal resolution, $k \in \{1, \ldots, 10\}$ indexes the shapelet within channel $\mathrm{c} \in \{1,2\}$. These shapelets are designed to capture class-discriminative waveform structures within localized temporal contexts. For each node $\mathrm{x_j} \in \mathbb{R}^{2\times W}$, which represents a dual-channel time window of size $W = 500$, the model computes the minimum squared Euclidean distance between every shapelet and all possible subsequences of length L within the corresponding channel. This operation is defined as:

$$d_{j,k}^{(c)} = \min_{\tau \in [), W-L]} \left\| x_j^{(c)} [\tau : \tau + \mathrm{L}] - s_k^{(c)} \right\|_2^2 \quad (4)$$

where $d_{j,k}^{(c)}$ denotes the best-matching distance between the k -th shapelet of channel c and the input node's signal. By performing this computation independently for both channels, the model generates 10 distances per modality, which are concatenated to form a node-level feature vector $\mathrm{z_j} \in \mathbb{R}^{20}$. This vector encapsulates how well each node's signal content aligns with the learned temporal patterns across modalities, effectively projecting each raw node signal into a discriminative shapelet-based feature space. This transformation replaces the raw node input with a modality-aware representation that is both compact and explainable, while also being robust to local temporal distortions such as phase shifts and misalignments. Importantly, because this process is applied independently to each node in the temporal graph, it enables localized pattern recognition within each time window while preserving channel-specific information. The shapelet parameters are learned jointly with the GAT in an end-to-end manner, ensuring that the extracted features are directly optimized for the downstream classification objective. Furthermore, the architecture includes an auxiliary computation routine that enables channel-specific averaging and masking, allowing for the quantitative analysis of the relative contributions of acoustic and optical modalities to the learned representations.

4.3 Graph neural network architecture/ graph-level embedding

Once the shapelet-based transformation is applied to each node in the temporal graph, the resulting feature vectors serve as input to a stack of GAT layers designed to model local temporal dependencies and enhance contextual understanding across the graph. GATs employ a self-attention mechanism that enables the network to dynamically weigh the importance of neighboring nodes during message passing. Specifically, for a node i and one of its neighbors j, an attention coefficient e_{ij} is computed as:

$$e_{ij} = \mathrm{LeakyReLU}\left(a^{\top}\left[Wh_i \parallel Wh_j\right]\right) \tag{5}$$

where W is a shared linear transformation, a is a learnable attention vector, and $\parallel$ denotes vector concatenation. These raw attention scores are normalized across the neighborhood of node i using the softmax function:

$$\alpha_{ij} = \frac{\exp\left(e_{ij}\right)}{\sum_{k \in \mathcal{N}(i)} \exp\left(e_{ik}\right)} \tag{6}$$

The updated node representation is then computed as a weighted sum of the transformed neighbor features, followed by a non-linear activation σ:

$$h_i{\prime} = \sigma\left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} W h_j\right) \tag{7}$$

This formulation allows the model to focus on salient temporal transitions while suppressing less informative connections, thereby enriching each node's feature representation with context-aware information. In our architecture, the shapelet-encoded graph is processed through a three-layer GAT network. The first two layers utilize multi-head attention (with two heads) to enhance stability and capture diverse relational patterns, followed by batch normalization and ELU non-linearities to improve convergence. The third layer employs a single-head GAT, also followed by batch normalization and ELU activation. Dropout is applied between layers to prevent overfitting and improve generalization. This deep GAT stack allows progressive refinement of node embeddings by incorporating temporal context from neighboring windows at multiple levels. To obtain a sample-level representation, the node embeddings $h_i^{(L)}$ from the final GAT layer are aggregated using global mean pooling:

$$z = \frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} h_i{}^{(L)} \tag{8}$$

The resulting graph-level embedding $z \in \mathbb{R}^{\mathrm{d}}$ captures both localized waveform characteristics (via shapelet extraction) and global temporal context (via attention-based aggregation). This embedding is passed through a two-layer fully connected classifier with ReLU activation and dropout, producing class logits $\hat{y} \in \mathbb{R}^{\mathrm{C}}$, where C is the number of target classes. The model is trained in a supervised fashion using the categorical cross-entropy loss. This loss encourages the model to assign high probability to the correct class while penalizing incorrect predictions proportionally. All components of the pipeline, including shapelet extraction, attention-based message passing, and classification, are optimized jointly in an end-to-end manner to minimize this objective and maximize predictive accuracy.

4.4 Multimodal shapelet-GAT classifier

The proposed model comprises three primary parameter groups: the shapelet extraction module, the GAT layers, and the classifier head. The shapelet extractor includes $2KL$ trainable parameters, where $K = 10$ shapelets are learned per channel and each shapelet has a length $L = 50$, resulting in a total of 1,000 learnable weights. Each GAT layer contributes parameters based on its configuration: with H attention heads, input dimensionality d_{in}, and output dimensionality d_{out}, the total number of parameters per layer is $H \cdot (d_{in} \cdot d_{out} + 2d)$, accounting for both the projection matrices and the learnable attention vectors. In our implementation, the GAT stack consists of three layers, where the first two employ multi-head attention (with two heads each), followed by a single-head final layer. This configuration yields only a few thousand trainable parameters, sufficient to model temporal dependencies and cross-window interactions while keeping the model lightweight. The classifier head comprises two fully connected layers: the first projects $d \rightarrow d$, and the second maps $d \rightarrow C$, adding $d^2 + dC$ parameters, where C denotes the number of output classes. Despite its compact size, the model retains strong expressiveness by combining low-dimensional, modality-specific shapelet features with graph-based temporal reasoning. Training is performed end-to-end using the Adam optimizer with a learning rate of 0.005. Each time-series sample is first converted into a graph structure through windowing and edge construction (as described in Sect. 3). The resulting graphs were partitioned using a stratified window-level split, with 65% used for training, 15% for validation, and 20% retained as a held-out test set. Stratification preserved class balance across the five nominal CuCrZr composition classes. The validation set was used for convergence monitoring and checkpoint selection, while the held-out test set was reserved exclusively for final performance evaluation. Because all samples originate from a single graded cube, this split represents a controlled within-build evaluation rather than a cross-build or cross-machine generalization test. The model was trained end-to-end for 300 epochs using categorical cross-entropy loss and the Adam optimizer. Batch normalization and dropout of 0.1 were applied after each GAT layer to improve training stability and reduce overfitting. The checkpoint with the highest validation accuracy was selected for final testing, and held-out test accuracy was used as the primary evaluation metric. Training loss, validation accuracy, and per-epoch wall-clock time were logged throughout training. All experiments were conducted with fixed random seeds for reproducibility and implemented using PyTorch and PyTorch Geometric on GPU-accelerated hardware.

4.5 Multi-material composition monitoring

4.5.1 Composition-sensitive process signatures in multi-material PBF-LB

Figure 12 illustrates the learning dynamics of the proposed Shapelet–GAT framework, depicted through two complementary metrics: training loss with standard deviation (left) and validation accuracy (right) over 300 training epochs. A sharp decline in training loss during the initial 50 epochs reflects rapid convergence, driven by the joint encoding of local discriminative motifs (via shapelets) and global relational context (via attention-based message passing). As training advances, the loss curve enters a smooth, asymptotic regime, suggesting convergence towards a stable local optimum. The narrowing of the variance band indicates increasing batch-wise stability and reduced gradient fluctuations, highlighting a well-behaved and generalizable optimization process. In parallel, the validation accuracy rises steadily, exceeding 92% by epoch 200, and remains consistently high thereafter with minimal fluctuations. The absence of divergence between training and validation curves suggests effective fitting without clear evidence of overfitting within the stratified within-build evaluation. This convergence behavior can be attributed to architectural regularization strategies such as dropout, batch normalization, and weight decay, which mitigate overfitting while preserving learning capacity. Collectively, these results validate the effectiveness of the modality-aware graph formulation, where multiscale temporal abstraction and graph-based relational reasoning are optimized in a compact parameter space. The observed learning dynamics affirm the robustness and scalability of the proposed pipeline for multimodal time-series classification under real-world variability Table 5.

Table 6 evaluates the final model's classification performance using a normalized confusion matrix across five copper composition classes. Remarkably, despite having only ~4,000 trainable parameters, the model achieves a classification accuracy of up to 92%, demonstrating a highly efficient yet expressive architecture. The pronounced diagonal dominance in the confusion matrix indicates strong precision and recall, with minimal misclassifications. Most errors are concentrated between adjacent compositions such

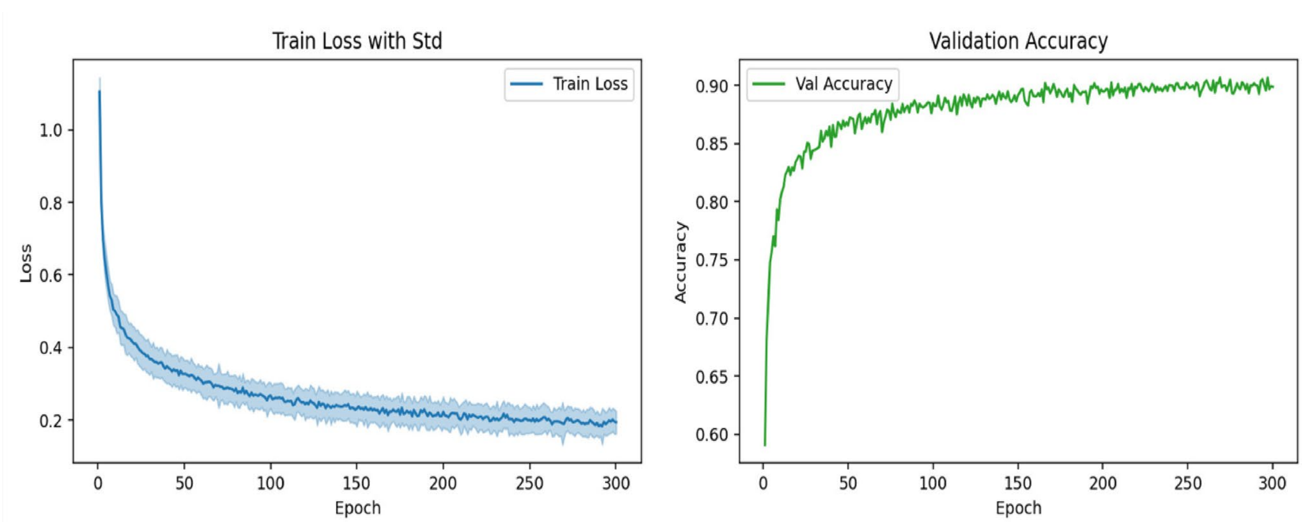

Fig. 12 Learning dynamics of the shapelet–GAT framework: Training loss and validation accuracy over 300 epochs

Table 5 Training parameters on the GAT models trained

Training parameters	GAT
Type of analysis	Sparse representation
Solver name	Adam
Learning rate	0.005
Training epochs	300
Batch size	256
Dropout	10%
Loss	Cross entropy
Shuffle	Every-epoch
Training set	65%
Testing and validation set	20% / 15%
Trainable parameter	3753

Table 6 Class-wise prediction accuracy of the proposed shapelet–GAT model using normalized confusion matrix

Classification (accuracy [%])	Ground truth				
	20%-Cu	40%-Cu	60%-Cu	80%-Cu	100%-Cu
20%-Cu	**87.2**	12.1	0.3	0.4	0.0
40%-Cu	6.5	**86.5**	5.5	1.5	0.0
60%-Cu	0.4	4.0	**90.7**	4.9	0.0
80%-Cu	0.5	0.5	4.9	**94.1**	0.0
100%-Cu	0.0	0.0	0.0	0.3	**99.7**

as 40%-Cu and 60%-Cu reflecting underlying physical and spectral similarities within these transition regions. Notably, the 100%-Cu class shows near-perfect classification, highlighting the model's ability to detect composition-specific features that are more distinctive at the distributional extremes. These results show that the learned shapelet and graph-attention representations capture class-discriminative dynamics across both sensing modalities. The model separates the five Cu-containing compositions most reliably at the compositional extremes, with modest confusion mainly between adjacent intermediate classes. Since all data originate from a single graded cube, the reported accuracy should be interpreted as within-build performance under a stratified training, validation, and test split, rather than evidence of cross-build or cross-machine generalization.

The $K = 10$ configuration discussed previously was fixed based on the assessment of the shapelet representation, as it achieved the highest accuracy while maintaining model compactness. To examine this effect, a limited sensitivity analysis was performed by varying the number of learnable shapelets per modality from $K = 2$ to $K = 16$, while keeping all other training and architectural hyperparameters constant. As shown in Fig. 13, classification accuracy increased markedly as K was increased from 2 to 10, improving from 79 to 92%. This indicates that very small shapelet banks under-represent the diversity of composition-sensitive temporal motifs present in the AE and OE signals. Since the number of learnable shapelets directly affects the parameter count, reducing K decreases model size but limits representational capacity, whereas increasing K provides additional temporal templates at the cost of higher model complexity. Further increasing K from 12 to 16 did not improve performance and instead reduced accuracy, reaching 88% at $K = 16$. Since only the number of shapelets was varied in this analysis, further optimization of related hyperparameters, such as shapelet length, graph connectivity, and attention-head configuration, may provide additional performance gains and remains an avenue for future work.

Saliency analysis was used as a post-hoc explanation tool to examine the internal decision logic of the proposed model. By quantifying the gradient-based sensitivity of the output with respect to input features, saliency maps identify the signal regions and sensing modalities that most strongly influence class-specific predictions. This not only facilitates

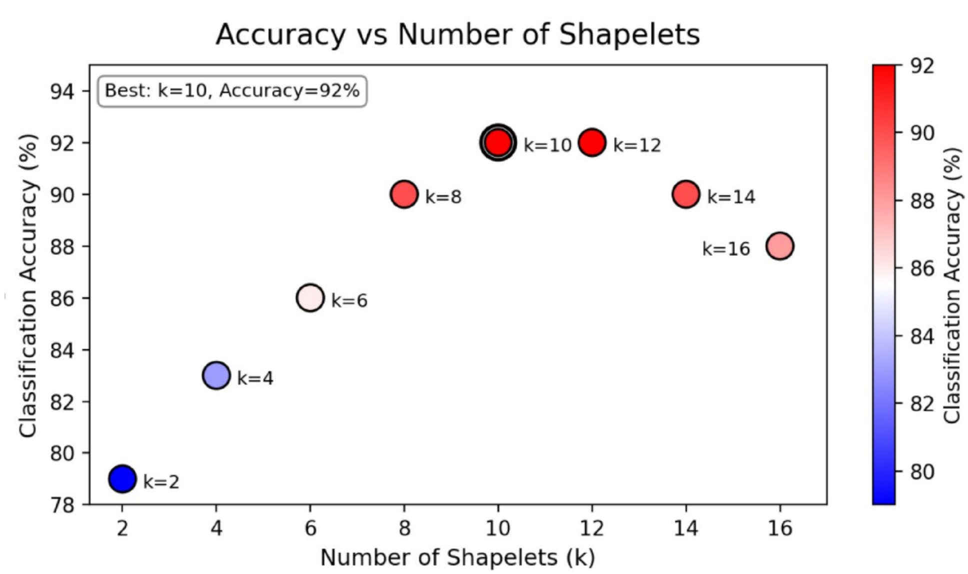

Fig. 13 Sensitivity of classification accuracy to the number of learnable shapelets per modality, varied from $K = 2$ to $K = 16$

understanding of the model's behavior. Let $\mathrm{x} \in \mathbb{R}^{N \times 2 \times T}$ denote the input tensor comprising N signal segments (nodes), each with two modalities—over T time steps. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ represent the multimodal graph constructed from these segments. During inference, we compute the gradient-based saliency $\nabla_{\mathrm{x}}\mathcal{L}$, where $\mathcal{L}$ is the classification loss. The absolute gradient with respect to each input channel is interpreted as its local importance:

$$S_{AE} = \sum_{t=1}^{T} \left| \frac{\partial \mathcal{L}}{\partial x_{:,0,t}} \right|, S_{Opt} = \sum_{t=1}^{T} \left| \frac{\partial \mathcal{L}}{\partial x_{:,1,t}} \right|, \tag{9}$$

These saliency scores are aggregated across nodes within each graph (i.e., sample) using a global pooling operation (e.g., global max-pooling):

$$\overline{S}_{AE}^{(i)} = Pool\left(S_{AE}^{(i)}\right), \overline{S}_{Opt}^{(i)} = Pool\left(S_{Opt}^{(i)}\right) \tag{10}$$

For each class c, we compute the mean saliency contribution per modality:

$$\mu_{AE}^{(c)} = \frac{1}{|D_c|} \sum_{\mathrm{i} \in D_c} \overline{S}_{AE}^{(i)}, \mu_{Opt}^{(c)} = \frac{1}{|D_c|} \sum_{\mathrm{i} \in D_c} \overline{S}_{Opt}^{(i)} \tag{11}$$

where D_c is the set of correctly classified samples belonging to class c. Finally, the percentage contribution of each modality is computed as:

$$AE\%^{(c)} = 100 \times \frac{\mu_{AE}^{(c)}}{\mu_{AE}^{(c)} + \mu_{Opt}^{(c)}}, Optical\%^{(c)} = 100 \times \frac{\mu_{Opt}^{(c)}}{\mu_{AE}^{(c)} + \mu_{Opt}^{(c)}} \tag{12}$$

Standard deviations σ^c are computed similarly to represent the intra-class variance in modality contribution. Figure 14 presents a quantitative analysis of modality-wise contributions to the classification process, computed using shapelet-level saliency scores aggregated across test samples. The bar plot compares the relative contributions of AE (CH1) and OE (CH2) across five copper composition classes. The results reveal a clear dominance of OE signals in the model's decision-making, especially for low to high copper contents (20%–80%). For instance, optical contributions exceed 60% for Classes 1–4 and peak at 64.9% for 40%-Cu. In contrast, AE shows relatively lower contributions across all classes, with a slight increase in 100%-Cu (40.7%), suggesting some resurgence of relevance for acoustic cues at high Cu concentration. This trend suggests that the model learns to adaptively prioritize sensing modalities based on their discriminative utility across the composition spectrum. The shifting contribution balance reflects not only changes in material reflectivity and emissivity but also the relative signal-to-noise characteristics captured by each sensor. Moreover, the low standard deviations across samples within each class indicate consistent modality usage, reinforcing the robustness of the model's attribution behavior. Importantly, this class-conditional, modality-aware attribution represents a key finding that has rarely been quantified or explicitly reported in prior sensor-fusion studies for additive manufacturing, moving beyond "fusion improves accuracy" to reveal how and when each modality drives the decision. These findings demonstrate that the model does not treat the two channels uniformly but instead performs class-conditional, modality-aware feature selection, which enhances interpretability and provides actionable guidance for optimizing sensor fusion and instrumentation strategies in multimodal process monitoring.

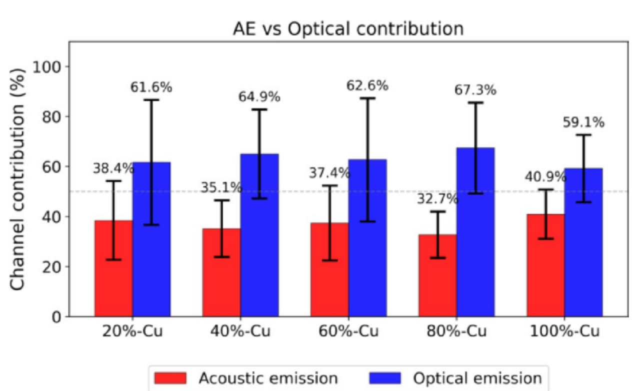

Fig. 14 Modality contribution (%) across classes based on saliency

Figure 15 illustrates the progression of each learnable shapelet from its initial random initialization to its final optimized form following end-to-end training. This evolution reflects the core capability of the shapelet extractor module to adaptively encode discriminative temporal structures from modality-specific dynamics. For Channel 1 (AE), the optimized shapelets consistently evolve towards high-frequency, burst-like motifs, indicating transient structural events or phase instabilities. In contrast, Channel 2 (OE) shapelets gravitate toward smoother, low-frequency trends, capturing gradual reflectance shifts likely related to thermal diffusion or surface morphology. This contrast between modalities exemplifies how distinct sensor modalities encode complementary information at differing temporal scales. The substantial divergence between initial and trained shapelet profiles confirms the effectiveness of gradient-based learning in driving task-specific temporal abstraction. Importantly, the diversity in final shapelet morphologies suggests that the model avoids mode collapse and preserves a rich representational capacity. Such evolution validates the utility of shapelet learning in uncovering latent motifs that would otherwise be obfuscated in raw time-domain representations.

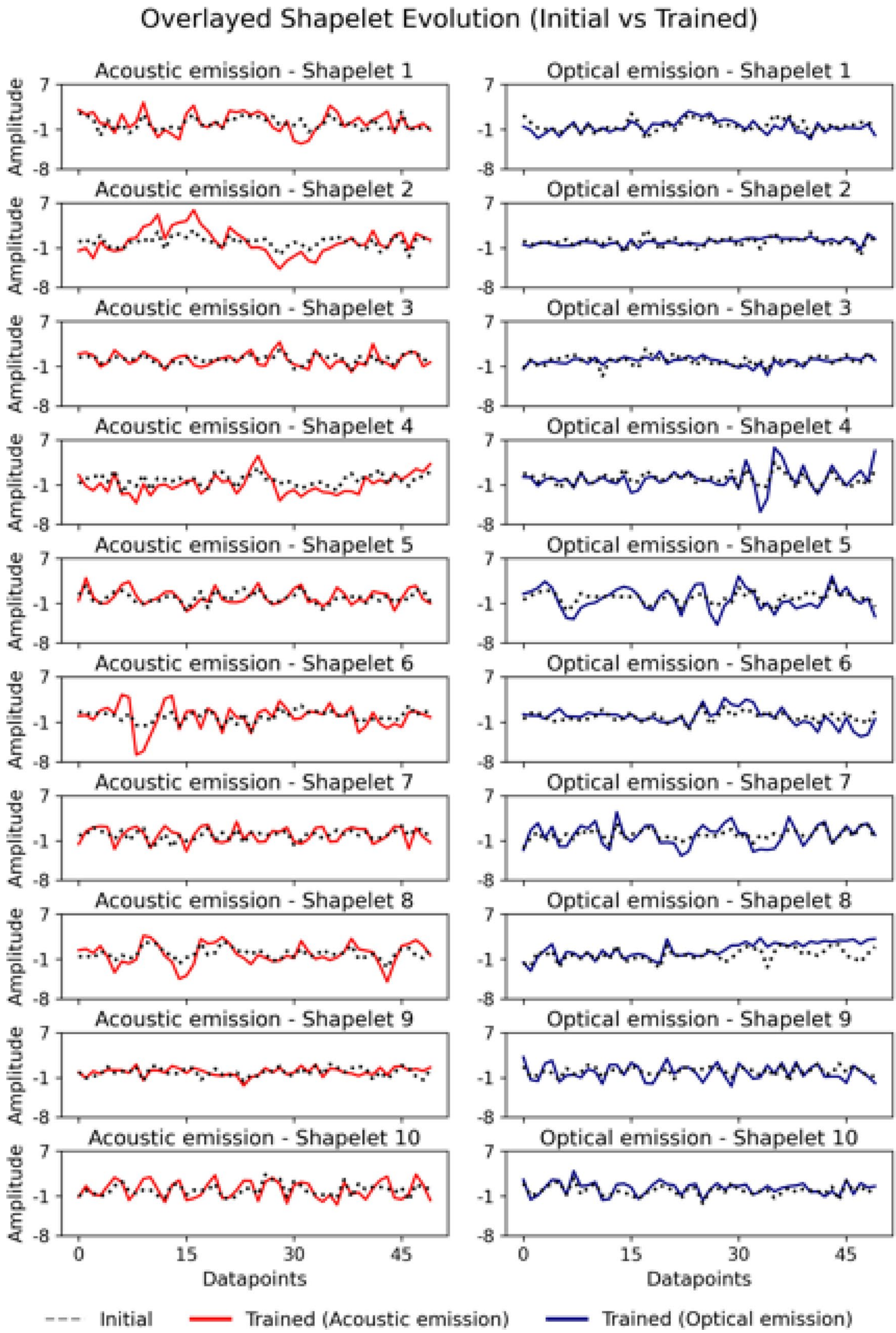

Fig. 15 Evolution of shapelets for channel 1 (red) and channel 2 (blue), overlaid with initialization

Figure 16 provides a fine-grained, class-wise visualization of node-wise shapelet activation maps across both sensing modalities: AE (left) and OE (right). Each heatmap represents the normalized distance between a learned shapelet and the corresponding signal window (node), where darker violet indicates stronger activation (i.e., lower distance between shapelet and segment). The columns represent nodes (i.e., sliding time windows), while rows represent different shapelets. In the AE maps, activations are relatively diffuse and low in intensity across all compositions, suggesting that the shapelets in Channel 1 are either weakly discriminative or capture more global, low-variance temporal features. For example, shapelet 1 exhibits near-uniform activation across many nodes and compositions, indicative of a structural motif that recurs regardless of class. This uniformity could reflect baseline acoustic signatures that are less composition-specific. Conversely, the OE channel reveals more sharply localized and heterogeneous activations, particularly in higher Cu compositions (60%, 80%, and 100%). Shapelets such as 1, 7, and 8 exhibit strong activation over specific node ranges (e.g., nodes 5–15), suggesting that optical motifs are temporally sparse but highly informative. These patterns may correspond to transient melt pool events or changes in reflectivity that emerge at particular process phases and are accentuated in high-Cu alloys. This interpretation is consistent with the transition from deeper keyhole/transition-mode melt pools at lower CuCrZr fractions to shallower conduction-mode melt pools at higher CuCrZr fractions. Therefore, the learned activation maps should not be interpreted as arbitrary neural features alone; they provide temporally localized indicators of how the dominant sensing mechanism shifts from mechanically driven acoustic responses toward optically driven reflectivity-sensitive responses across the compositional gradient. This detailed spatial–temporal activation mapping complements the earlier saliency attribution analysis (Fig. 15) by revealing not just how much a modality contributes, but where and when within a sample the most discriminative waveform patterns occur. The pronounced and localized activations in Channel 2 reinforce the conclusion that OE plays a critical role in classifying Cu-rich compositions. Overall, this map-based inspection illustrates the modality-aware and temporally adaptive learning behavior of the proposed Shapelet–GAT framework, offering interpretability through the ability to trace predictions back to salient motifs in specific time segments.

Figure 17 shows the normalized node-wise saliency across 19 temporal segments for each composition class. Higher saliency at nodes 1 and 19 indicates that the model relies more strongly on the beginning and end of each scan-track-aligned signal window. This is physically plausible, as these boundary regions are likely to contain laser-on and laser-off transients, melt-pool initiation and termination, and rapid changes in thermal and solidification conditions. In contrast, the central nodes show lower and more uniform saliency, suggesting a smaller contribution to classification. The trend is broadly consistent across classes, with slightly stronger variation for 80%-Cu and 100%-Cu. Although the learned shapelets are temporal motifs rather than explicit spectral filters, their morphology is consistent with the exploratory spectral analysis: AE contributes mainly through transient, burst-like waveform content, while OE captures smoother intensity variations and reflectivity-driven signatures that become more prominent at higher CuCrZr fractions. Overall, the node-wise saliency confirms that the Shapelet–GAT model focuses on physically meaningful and temporally informative signal regions.

Figure 18 visualizes the learned graph-level embeddings using a 2D *t*-SNE projection, with samples color-coded by Cu composition. The extreme classes, 20%-Cu and 100%-Cu, form compact and clearly separated clusters, indicating that the model captures distinct signal morphologies at the compositional boundaries. In contrast, intermediate compositions, particularly 40%-Cu and 60%-Cu, show greater overlap, consistent with their closer physical similarity and the modest confusion observed in the classification matrix. The gradual transition among clusters suggests that the latent space preserves composition-dependent structure rather than only separating discrete classes. Overall, the *t*-SNE visualization supports that the learned graph-level embeddings preserve composition-dependent structure, with clearer separation at the compositional extremes and partial overlap between intermediate Cu-containing classes.

5 Discussion

While the preceding sections demonstrated the effectiveness of the multimodal Shapelet–GAT framework, it is equally important to evaluate its generalizability and adaptability under constrained sensing conditions. To this end, the architecture was tested in unimodal configurations, using only either AE or OE data. This evaluation isolates the contribution of each sensor stream and tests the robustness of the framework when operating with limited input data. As shown in Table 7, the unimodal acoustic and optical configurations achieve classification accuracies of ≈82.0% and ≈81%, respectively. The unimodal results show that both AE and OE contain composition-sensitive information, but neither modality alone reaches the performance of the dual-channel model. The variations observed in AE and OE arise from the coupled effects of composition changes and melting mode transitions. The shift from keyhole to conduction mode modifies melt pool depth and stability, influencing the

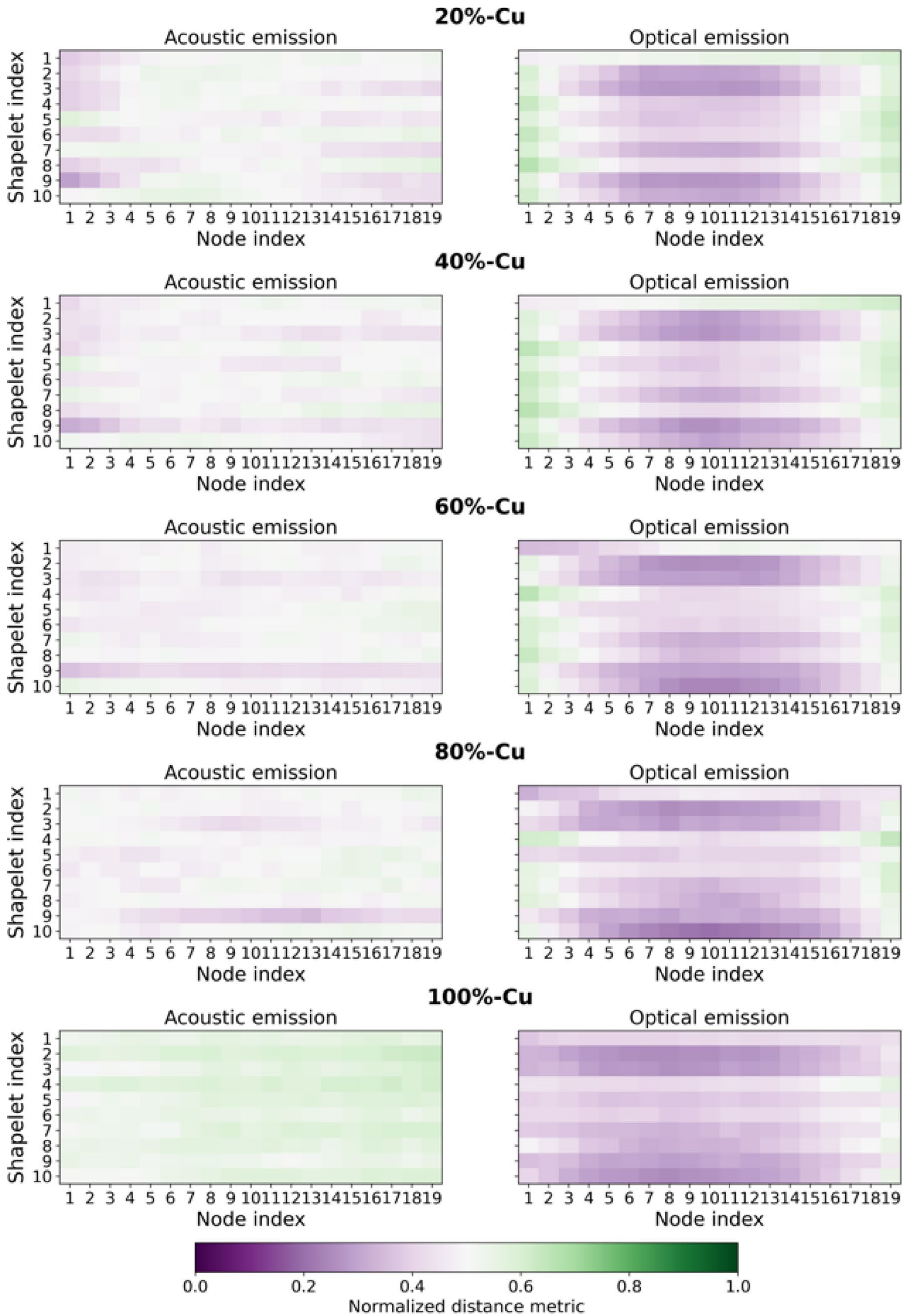

Fig. 16 Class-wise shapelet activation maps across time nodes and sensing modalities

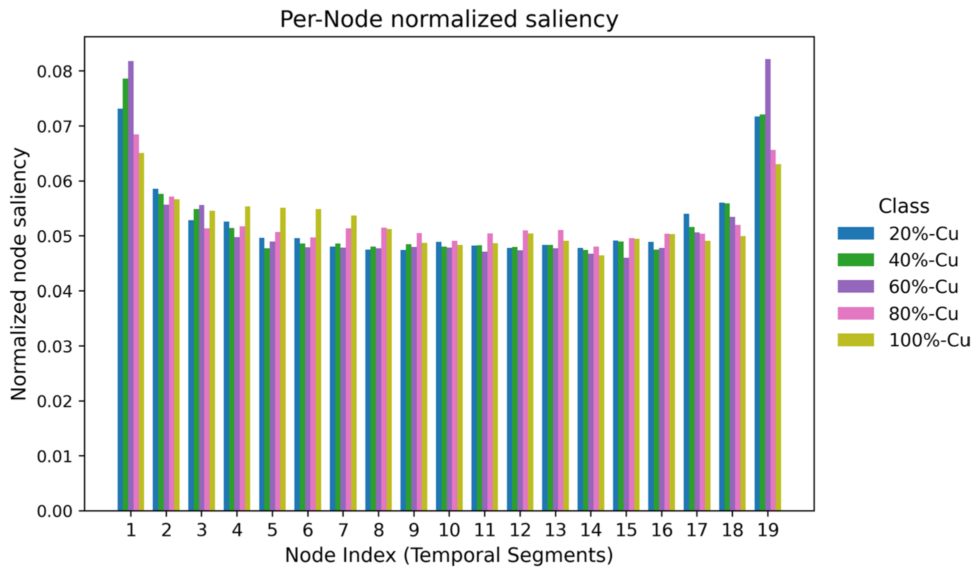

Fig. 17 Node-wise normalized saliency across temporal segments and composition classes

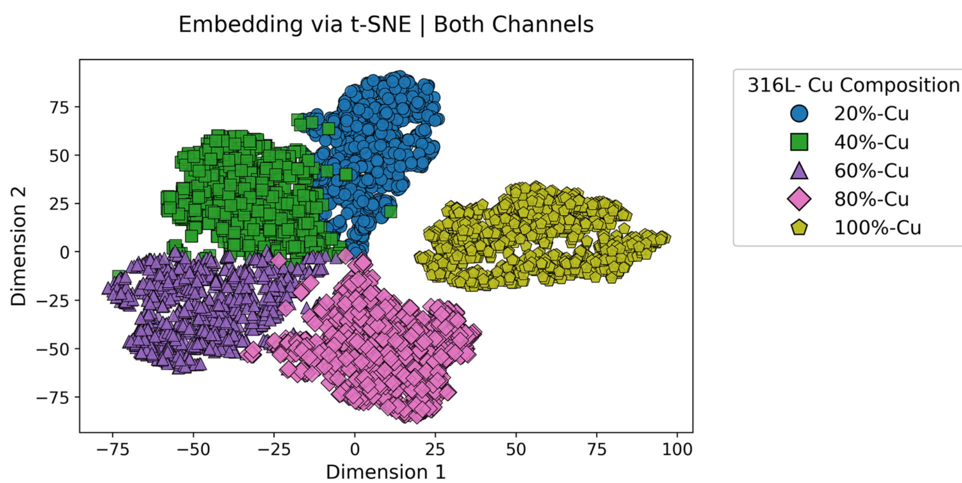

Fig. 18 Latent space structure revealed by *t*-SNE: class-coherent and composition-aware clusters

Table 7 Performance comparison of shapelet–GAT under unimodal configurations

Classification (accuracy [%])	Ground truth				
	20%-Cu	40%-Cu	60%-Cu	80%-Cu	100%-Cu
20%-Cu	***75.5***	***19.0***	***3.4***	***2.1***	***0.0***
	76.5	18.7	1.8	3.0	0.1
40%-Cu	***20.0***	***68.8***	***9.7***	***1.5***	***0.0***
	13.7	65.1	19.0	2.2	0.0
60%-Cu	***0.9***	***7.2***	***84.2***	***7.6***	***0.1***
	2.3	13.2	77.8	6.7	0.0
80%-Cu	***0.9***	***2.1***	***12.5***	***84.5***	***0.0***
	1.5	1.7	12.7	84.2	0.0
100%-Cu	***0.0***	***0.0***	***0.0***	***0.0***	***100.0***
	0.0	0.0	0.0	0.0	100.0

Each cell contains classification accuracies for the unimodal Shapelet–GAT architecture, with ***bold italicized*** values representing airborne acoustic-only results and plain values representing optical-only results

frequency content and amplitude of AE signals. In parallel, increasing CuCrZr content changes optical absorptivity and reflectivity, producing systematic variations in OE intensity and spectral response. Thus, AE mainly reflects melt pool dynamics and mechanical stability, while OE is more sensitive to composition related changes in reflectivity and absorptivity. The improved dual modality performance is therefore physically consistent with the complementary nature of the two sensing channels.

To further evaluate whether the proposed Shapelet–GAT architecture is justified beyond unimodal sensor ablation, additional benchmarking was performed against representative end-to-end time-series and graph-learning baselines, as summarized in Table 8. The 1D-CNN baseline processes the dual-channel AE/OE signal using three stacked one-dimensional convolutional layers with 32, 64, and 128 filters, followed by batch normalization, ReLU activation, adaptive average pooling, and a compact fully connected classifier. The CNN–LSTM baseline uses two convolutional layers

Table 8 Single-run benchmark comparison of representative end-to-end time-series and graph-learning baselines

Model	Accuracy (%)	F1-score (%)	Params	Model footprint (KB)	FLOPs	Train time (s)	Amortized GPU latency (ms/sample)	Peak GPU memory (MB)
Shapelet-GAT	91.64	91.7	3.8k	14.66	25 M	550.2	0.031	1200
1D-CNN	91.53	91.34	44.5k	173.89	30.33M	386.82	0.029	654.4
CNN-LSTM	91.40	91.45	86.1k	336.39	25.17M	231.45	0.020	1125.3
TCN	76.60	76.39	115.9k	452.64	538.92M	5003.44	0.298	9021.8
Transformer	83.24	83.08	82.1k	320.52	4.02M	188.75	0.015	226.5
GAT without shapelets	78.63	78.69	22.8k	88.97	0.39M	343.82	0.046	107.1
GCN without shapelets	76.65	76.83	22.6k	88.35	0.39M	297.60	0.039	66.9

with 32 and 64 filters to extract local temporal features, followed by a bidirectional LSTM with 64 hidden units to model longer-range sequential dependencies. The TCN baseline consists of three dilated residual convolutional blocks with 32, 64, and 128 channels and dilation factors of 1, 2, and 4, enabling multi-scale temporal feature extraction. The Transformer baseline first projects the 5000-point dual-channel input into 100 temporal tokens using a stride-50 convolutional projection, followed by a two-layer Transformer encoder with 64-dimensional embeddings and four attention heads. To isolate the contribution of the shapelet module, GAT and GCN baselines without shapelet extraction were also implemented, where each 500-point dual-channel node is flattened, projected to a 20-dimensional representation, and passed through three graph-attention or graph-convolution layers. These benchmark models should be interpreted as representative baseline implementations rather than exhaustively optimized architectures; further hyperparameter tuning, pruning, quantization, or architecture-specific optimization could improve their performance or efficiency. Nevertheless, they provide a practical comparison across conventional convolutional, recurrent, temporal-convolutional, attention-based, and graph-based learning strategies. For a fair comparison, all results reported in Table 8 were generated using the same stratified 65/15/20 train/validation/test split described in Sect. 5.

The term lightweight is used here specifically to denote the compact trainable-parameter footprint of the model, rather than lower peak runtime memory or lower inference latency reported in Table 8. Model footprint was calculated as the number of trainable parameters multiplied by 4 bytes, assuming float32 precision. Peak GPU memory and amortized GPU latency were measured on an NVIDIA RTX PRO 4000 Blackwell GPU using a batch size of 256. The reported latency values represent total batched inference time divided by 256 and should therefore not be interpreted as isolated batch-size-one inference latency. Although the proposed Shapelet–GAT has a small parameter footprint, its fully connected graph construction, shapelet-distance computation, and graph-attention operations increase FLOPs, activation storage, and runtime memory. Therefore, hardware-specific and library-level optimization will be required to fully exploit the model's compact parameter footprint and to assess its practical runtime memory, latency, and deployment efficiency under resource-constrained manufacturing environments.

As shown in Table 8, the Shapelet–GAT model achieved 91.64% accuracy with only 3.8k trainable parameters, while the 1D-CNN and CNN–LSTM baselines achieved closely comparable accuracies of 91.53% and 91.40%, respectively, but required substantially larger parameter counts of 44.5k and 86.1k. Since the accuracy difference among these three models is only 0.24 percentage points in a single-run benchmark, the results should not be interpreted as a statistically definitive ranking or as evidence of clear accuracy superiority. Rather, the value of the proposed framework lies in achieving comparable within-build classification performance with a compact trainable-parameter footprint, while providing interpretable access to temporal shapelets, modality-wise saliency, node-wise importance, and sensor-contribution trends. The near-identical accuracies of the strongest dual-modality neural models also suggest that the present five-class, single-cube task may be close to saturation for high-performing end-to-end architectures under the current window-level evaluation setting. Therefore, Table 8 should be viewed primarily as a representative benchmark showing that multiple neural architectures can extract composition-sensitive AE/OE signatures, rather than as a definitive comparison of architectural superiority. Extending equivalent size–accuracy sweeps to the baseline architectures would require a broader capacity-matched benchmarking study and is therefore treated as future work rather than the main objective of the present manuscript.

In contrast, the graph baselines without shapelet extraction performed substantially worse, with GAT without shapelets reaching 78.63% accuracy and GCN without shapelets reaching 76.65%. This indicates that graph-based temporal reasoning alone is insufficient, and that the learnable shapelet representation provides an important discriminative temporal abstraction. The benchmark also reports floating point operations (FLOPs), inference latency, memory footprint, and training time, providing a practical comparison of predictive performance and computational cost. Classical models such as Random Forest, support vector machines, and eXtreme Gradient Boosting were not used as primary baselines because they require manually engineered time- and frequency-domain features, whereas the proposed framework and selected neural baselines operate directly on synchronized AE and OE signal windows.

The proposed Shapelet–GAT model provides a compact architecture with approximately 4,000 trainable parameters while maintaining classification performance comparable to larger neural baselines. The benchmark results should be interpreted as a practical comparison of computational complexity rather than as a full embedded deployment study. Dedicated hardware validation on embedded processors or machine controller hardware remains necessary before making definitive claims regarding real-time edge implementation. The added value of the framework is not limited to classification accuracy. Compared with the earlier contrastive learning approach [67], which focused on learning composition separable embeddings from fused AE and OE signals, the Shapelet–GAT model provides explicit access to the temporal motifs, sensor modality contributions, and graph node regions that drive the prediction. This is important for multi-material PBF-LB, where process signatures arise from coupled effects of composition, reflectivity, melt pool morphology, and melting regime. The interpretability analyses show that the two sensing channels are used differently across the compositional gradient, with OE becoming increasingly influential at higher CuCrZr concentrations, consistent with increased reflectivity near the laser wavelength. Overall, the learnable shapelet representation provides a compact and interpretable mechanism for encoding discriminative temporal motifs, supporting the Shapelet–GAT model as an interpretable sensor fusion framework for composition sensitive monitoring in multi-material PBF-LB.

6 Conclusion

This study presents a compact and interpretable Shapelet–GAT framework for decoding composition-dependent AE and OE signatures during 316L–CuCrZr multi-material PBF-LB. By combining learnable shapelets with graph-based attention, the model captures localized waveform patterns and temporal relationships from synchronized dual-modality sensor data. With approximately 4,000 trainable parameters, it achieves classification accuracy comparable to larger neural baselines while providing interpretable access to modality-specific temporal motifs and sensor-contribution trends. Overall, the framework provides an interpretable graph–shapelet alternative to conventional deep learning approaches for multimodal AM process monitoring. Three key findings underscore the significance of this framework:

1. The multimodal Shapelet–GAT model achieved up to 92% classification accuracy, outperforming unimodal configurations and achieving performance comparable to larger neural baseline models, while maintaining a substantially smaller trainable-parameter footprint.
2. Saliency analysis and shapelet activation maps provided deeper, modality-specific insight into the decision logic, showing that back-reflected OE at the laser wavelength becomes increasingly dominant at higher Cu concentrations, consistent with copper's higher reflectivity at 1064 nm.
3. The learned latent embedding space, visualized via *t*-SNE, exhibited well-structured, composition-aware clusters with smooth transitions along the copper gradient—capturing both discrete class boundaries and continuous material variation.

These findings reinforce our previous observation that OE becomes increasingly informative in reflective CuCrZr-rich regions of multi-material PBF-LB [67]. However, the present validation is limited to a single graded build with a stratified window-level split for training, validation, and testing. Independent-build and layer-wise validation are therefore required to establish cross-build generalization. Accordingly, the framework represents a proof-of-concept for composition-sensitive signal decoding, with off-nominal validation required before extension to anomaly detection and closed-loop quality assurance. Overall, it provides a strong foundation for advancing data-driven monitoring in smart manufacturing. Future work will explore its extension to variable-length sequences [68], integration of additional sensing modalities (e.g., thermal imaging, spectroscopy, high-speed video), and cross-system generalization across different machines, build geometries, and alloy families. In addition, hardware-specific and library-level optimization will be required to fully exploit the model's compact parameter footprint and to assess its practical runtime memory, latency, and deployment efficiency under resource-constrained manufacturing environments. Further enhancements may also include embedding domain-aligned constraints into shapelet learning or attention mechanisms to improve robustness and support adaptive monitoring in complex, multi-material AM environments. The data and code for this study can be accessed in the following repository (https://gitlab.utu.fi/vpsora/Additive-Manufacturing-M ulti-Material-316l_CuCrZr-Sensor-Fusion-Modality-Aware-Lightweight-Graph-Network.git).

Author Contribution V.P. conceptualized the study, developed the methodology, performed the formal analysis, implemented the software, validated the results, prepared the visualizations, and wrote the original draft. A.B. contributed to the conceptualization, formal analysis, validation, and writing of the original draft. A.S. supervised the work and reviewed and edited the manuscript. C.L. contributed to the conceptualization and methodology, supervised the work, and reviewed and edited the manuscript. All authors reviewed and approved the final manuscript.

Funding Open Access funding provided by University of Turku (including Turku University Central Hospital). This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability For those interested, the data and codes for this study can be accessed in the following repository (https://gitlab.utu.fi/vpsora/Additive-Manufacturing-Multi-Material-316l_CuCrZr-Sensor-Fusion-Modality-Aware-Lightweight-Graph-Network.git).

Declarations

Conflict Interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. DebRoy T, Wei HL, Zuback JS, Mukherjee T, Elmer JW, Milewski JO, Beese AM, Wilson-Heid A, De A, Zhang W (2018) Additive manufacturing of metallic components–process, structure and properties. Prog Mater Sci 92:112–224
2. Herzog D, Seyda V, Wycisk E, Emmelmann C (2016) Additive manufacturing of metals. Acta Mater 117:371–392
3. Frazier WE (2014) Metal additive manufacturing: a review. J Mater Eng Perform 23:1917–1928
4. Blakey-Milner B, Gradl P, Snedden G, Brooks M, Pitot J, Lopez E, Leary M, Berto F, Du Plessis A (2021) Metal additive manufacturing in aerospace: a review. Mater Des 209:110008
5. Z. Mehdiyev, C. Felho, Metal additive manufacturing in automotive industry: A review of applications, advantages, and limitations, in: Materials Science Forum, Trans Tech Publ, 2023, pp. 49–62.
6. Sun C, Wang Y, McMurtrey MD, Jerred ND, Liou F, Li J (2021) Additive manufacturing for energy: a review. Appl Energy 282:116041
7. Li Y, Jiang D, Zhu R, Yang C, Wang L, Zhang L-C (2024) Revolutionizing medical implant fabrication: advances in additive manufacturing of biomedical metals. Int J Extrem Manuf 7:022002
8. Russell R, Wells D, Waller J, Poorganji B, Ott E, Nakagawa T, Sandoval H, Shamsaei N, Seifi M (2019) Qualification and certification of metal additive manufactured hardware for aerospace applications. Additive manufacturing for the aerospace industry. pp 33–66
9. Baig S, Jam A, Beretta S, Shao S, Shamsaei N (2025) Nondestructive detection of critical defects in additive manufacturing. Sci Rep 15:6740
10. Duarte VR, Rodrigues TA, Machado MA, Pragana JP, Pombinha P, Coutinho L, Silva CM, Miranda RM, Goodwin C, Huber DE (2021) Benchmarking of nondestructive testing for additive manufacturing. 3D Print Addit Manuf 8:263–270
11. L. Koester, H. Taheri, T. Bigelow, L. Bond, Nondestructive testing for metal parts fabricated using powder based additive manufacturing, Materials evaluation, 76 (2018).
12. Wang J, Zhu R, Liu Y, Zhang L (2023) Understanding melt pool characteristics in laser powder bed fusion: an overview of single- and multi-track melt pools for process optimization. Adv Powder Mater 2:100137
13. Miner JP, Ngo A, Gobert C, Reddy T, Lewandowski JJ, Rollett AD, Beuth J, Narra SP (2024) Impact of melt pool geometry variability on lack-of-fusion porosity and fatigue life in powder bed fusion-laser beam Ti–6Al–4V. Addit Manuf 95:104506
14. D. Guirguis, Machine learning-Enabled Multi-scale Process Monitoring and Development for Metal Additive Manufacturing, in, Carnegie Mellon University.
15. Chen L, Bi G, Yao X, Su J, Tan C, Feng W, Benakis M, Chew Y, Moon SK (2024) In-situ process monitoring and adaptive quality enhancement in laser additive manufacturing: a critical review. J Manuf Syst 74:527–574
16. McCann R, Obeidi MA, Hughes C, McCarthy É, Egan DS, Vijayaraghavan RK, Joshi AM, Garzon VA, Dowling DP, McNally PJ (2021) In-situ sensing, process monitoring and machine control in laser powder bed fusion: a review. Addit Manuf 45:102058
17. K.R. Gutknecht, In-situ laser powder bed fusion process monitoring and control, in, ETH Zurich, 2023.
18. Kavas B, Balta EC, Tucker MR, Krishnadas R, Rupenyan A, Lygeros J, Bambach M (2025) In-situ controller autotuning by Bayesian optimization for closed-loop feedback control of laser powder bed fusion process. Addit Manuf 99:104641
19. Cai Y, Xiong J, Chen H, Zhang G (2023) A review of in-situ monitoring and process control system in metal-based laser additive manufacturing. J Manuf Syst 70:309–326
20. C.Y. Yap, C.K. Chua, Z.L. Dong, Z.H. Liu, D.Q. Zhang, L.E. Loh, S.L. Sing, Review of selective laser melting: Materials and applications, Appl Phys Rev, 2 (2015).
21. Chowdhury S, Yadaiah N, Prakash C, Ramakrishna S, Dixit S, Gupta LR, Buddhi D (2022) Laser powder bed fusion: a state-of-the-art review of the technology, materials, properties & defects, and numerical modelling. J Mater Res Technol 20:2109–2172
22. Oliveira JP, LaLonde A, Ma J (2020) Processing parameters in laser powder bed fusion metal additive manufacturing. Mater Des 193:108762
23. Gordon JV, Narra SP, Cunningham RW, Liu H, Chen H, Suter RM, Beuth JL, Rollett AD (2020) Defect structure process maps for laser powder bed fusion additive manufacturing. Addit Manuf 36:101552
24. Stopka KS, Desrosiers A, Andreaco A, Sangid MD (2024) A methodology for the rapid qualification of additively manufactured

materials based on pore defect structures. Integr Mater Manuf Innov 13:335–359
25. Gunasegaram DR, Barnard AS, Matthews MJ, Jared BH, Andreaco AM, Bartsch K, Murphy AB (2024) Machine learning-assisted in-situ adaptive strategies for the control of defects and anomalies in metal additive manufacturing. Addit Manuf 81:104013
26. Everton SK, Hirsch M, Stravroulakis P, Leach RK, Clare AT (2016) Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing. Mater Des 95:431–445
27. Tan C, Li R, Su J, Du D, Du Y, Attard B, Chew Y, Zhang H, Lavernia EJ, Fautrelle Y (2023) Review on field assisted metal additive manufacturing. Int J Mach Tools Manuf 189:104032
28. Y. Sun, S. Gorgannejad, A. Martin, J. Nicolino, M. Strantza, J.-B. Forien, V. Thampy, S. Liu, P. Quan, C.J. Tassone, Direct mechanistic connection between acoustic signals and melt pool morphology during laser powder bed fusion, Appl Phys Letters, 125 (2024).
29. Forien J-B, Guss GM, Khairallah SA, Smith WL, DePond PJ, Matthews MJ, Calta NP (2023) Detecting missing struts in metallic micro-lattices using high speed melt pool thermal monitoring. Addit Manuf Lett 4:100112
30. Zhang Y, Fuh JY, Ye D, Hong GS (2019) In-situ monitoring of laser-based PBF via off-axis vision and image processing approaches. Addit Manuf 25:263–274
31. Zhang W, Ma H, Zhang Q, Fan S (2022) Prediction of powder bed thickness by spatter detection from coaxial optical images in selective laser melting of 316L stainless steel. Mater Des 213:110301
32. J. Williams, K. Snodderly, M. Kottman, D. Paredes, M. Jamshid, M. White, M. Seifi, D. Wells, E. Lanigan, M. James, Strategic Guide: Additive Manufacturing In-Situ Monitoring Technology Readiness, ASTM Int Additive Manuf Cent Excell, (2023).
33. Gaikwad A, Williams RJ, de Winton H, Bevans BD, Smoqi Z, Rao P, Hooper PA (2022) Multi phenomena melt pool sensor data fusion for enhanced process monitoring of laser powder bed fusion additive manufacturing. Mater Des 221:110919
34. Bevans B, Barrett C, Spears T, Gaikwad A, Riensche A, Smoqi Z, Halliday H, Rao P (2023) Heterogeneous sensor data fusion for multiscale, shape agnostic flaw detection in laser powder bed fusion additive manufacturing. Virt Phys Prototyp 18:e2196266
35. Tempelman JR, Wachtor AJ, Flynn EB, Depond PJ, Forien J-B, Guss GM, Calta NP, Matthews MJ (2022) Sensor fusion of pyrometry and acoustic measurements for localized keyhole pore identification in laser powder bed fusion. J Mater Process Technol 308:117656
36. Herzog T, Brandt M, Trinchi A, Sola A, Molotnikov A (2024) Process monitoring and machine learning for defect detection in laser-based metal additive manufacturing. J Intell Manuf 35:1407–1437
37. Pandiyan V, Masinelli G, Claire N, Le-Quang T, Hamidi-Nasab M, de Formanoir C, Esmaeilzadeh R, Goel S, Marone F, Logé R (2022) Deep learning-based monitoring of laser powder bed fusion process on variable time-scales using heterogeneous sensing and operando X-ray radiography guidance. Addit Manuf 58:103007
38. Pandiyan V, Drissi-Daoudi R, Shevchik S, Masinelli G, Le-Quang T, Logé R, Wasmer K (2021) Semi-supervised monitoring of laser powder bed fusion process based on acoustic emissions. Virt Phys Prototyp 16:481–497
39. Fu Y, Downey AR, Yuan L, Zhang T, Pratt A, Balogun Y (2022) Machine learning algorithms for defect detection in metal laser-based additive manufacturing: a review. J Manuf Process 75:693–710
40. Pandiyan V, Wróbel R, Richter RA, Leparoux M, Leinenbach C, Shevchik S (2024) Monitoring of laser powder bed fusion process by bridging dissimilar process maps using deep learning-based domain adaptation on acoustic emissions. Addit Manuf 80:103974
41. Gutknecht K, Haferkamp L, Cloots M, Wegener K (2020) Determining process stability of laser powder bed fusion using pyrometry. Procedia Cirp 95:127–132
42. Lapointe S, Guss G, Reese Z, Strantza M, Matthews MJ, Druzgalski CL (2022) Photodiode-based machine learning for optimization of laser powder bed fusion parameters in complex geometries. Addit Manuf 53:102687
43. Pandiyan V, Wróbel R, Richter RA, Leparoux M, Leinenbach C, Shevchik S (2023) Self-supervised Bayesian representation learning of acoustic emissions from laser powder bed fusion process for in-situ monitoring. Mater Des 235:112458
44. Oster S, Breese PP, Ulbricht A, Mohr G, Altenburg SJ (2024) A deep learning framework for defect prediction based on thermographic in-situ monitoring in laser powder bed fusion. J Intell Manuf 35:1687–1706
45. Snow Z, Diehl B, Reutzel EW, Nassar A (2021) Toward in-situ flaw detection in laser powder bed fusion additive manufacturing through layerwise imagery and machine learning. J Manuf Syst 59:12–26
46. Dharmadhikari S, Menon N, Basak A (2023) A reinforcement learning approach for process parameter optimization in additive manufacturing. Addit Manuf 71:103556
47. Ogoke F, Farimani AB (2021) Thermal control of laser powder bed fusion using deep reinforcement learning. Addit Manuf 46:102033
48. Gutknecht K, Cloots M, Sommerhuber R, Wegener K (2021) Mutual comparison of acoustic, pyrometric and thermographic laser powder bed fusion monitoring. Mater Des 210:110036
49. Zhang S, Zhang Z, Wang J, Huang J, Qin R, Qin H, Li Z, Wen G, Zhang Q, Chen X (2025) A graph neural network integrating physical prior knowledge for defect monitoring in laser powder bed fusion. J Manuf Process 153:516–530
50. A. Karkadakattil, Physics-informed machine learning framework for grain size prediction in laser-processed 316L stainless steel: integration of experimental microstructure and surrogate data, Canadian Metallurgical Quarterly, (2025) 1–14.
51. A. Karkadakattil, Physics-guided neural network framework for surface roughness prediction in additively manufactured metallic alloys, J Laser Appl, 38 (2026).
52. Pandiyan V, Wróbel R, Shevchik S, Leinenbach C (2026) Explainable zero-shot transfer learning for cross-domain in-situ acoustic monitoring in laser powder bed fusion process using learnable wavelet scattering. Mater Des 265:115888
53. Wei C, Li L (2021) Recent progress and scientific challenges in multi-material additive manufacturing via laser-based powder bed fusion. Virt Phys Prototyp 16:347–371
54. Mehrpouya M, Tuma D, Vaneker T, Afrasiabi M, Bambach M, Gibson I (2022) Multimaterial powder bed fusion techniques. Rapid Prototyp J 28:1–19
55. Ren C, Liang T, Jin S, Song Z, Dan X, Liu Q, Sun Y, Wang M, Du Y, Wang C (2025) Exceptional strength and antibacterial durability in hierarchically structured Cu-bearing 316L stainless steel through additive manufacturing. J Mater Res Technol. https://doi.org/10.1016/j.jmrt.2025.04.220
56. Griffis J, Shahed K, Meinert K, Yilmaz B, Lear M, Manogharan G (2025) Multi-material laser powder bed fusion: effects of build orientation on defects, material structure and mechanical properties. npj Adv Manuf 2:5
57. Sokkalingam R, Åsberg M, Krakhmalev P (2025) In-situ alloying of Cu in 316L stainless steel by PBF-LB: influence of laser power and rescanning strategy. J Mater Res Technol 35:6137–6146
58. Li X, Pan Z, Smolej L, Nadimpalli VK, Moshiri M (2024) Towards manufacturing intra-layer multi-material mould tools

with vertical interfaces using laser-based powder bed fusion. Mater Des 243:113056
59. Liu L, Wang D, Deng G, Han C, Yang Y, Chen J, Chen X, Liu Y, Bai Y (2022) Laser additive manufacturing of a 316L/CuSn10 multimaterial coaxial nozzle to alleviate spattering adhesion and burning effect in directed energy deposition. J Manuf Process 82:51–63
60. Gao Z, Yang Y, Wang L, Zhou B, Yan F (2022) Formation mechanism and control of solidification cracking in laser-welded joints of steel/copper dissimilar metals. Metals 12:1147
61. Sun X, Hao W, Geng G, Ma T, Li Y (2018) Solidification microstructure evolution of undercooled Cu-15 wt.% Fe alloy melt. Adv Mater Sci Eng. https://doi.org/10.1155/2018/6304518
62. Martendal CP, Esteves PDB, Deillon L, Malamud F, Jamili AM, Löffler JF, Bambach M (2024) Effects of beam shaping on copper-steel interfaces in multi-material laser beam powder bed fusion. J Mater Process Technol 327:118344
63. Zhang X, Li L, Liou F (2021) Additive manufacturing of stainless steel–copper functionally graded materials via Inconel 718 interlayer. J Mater Res Technol 15:2045–2058
64. Bobbio L, Bocklund B, Liu Z, Beese A, Otis R, Borgonia J, Dillon R, Shapiro A, McEnerney B (2018) Characterization of a functionally graded material of Ti-6Al-4V to 304L stainless steel with an intermediate V section. J Alloys Compd 742:1031–1036
65. A. Baganis, F. Malamud, J. Čapek, F.F. Klimashin, E. Polatidis, M. Busi, M. Šmíd, M. Jambor, J. Michler, M. Strobl, Phase formation and texture evolution in 316L-CuCrZr multi-material structures fabricated by laser powder bed fusion, Materials & Design, (2025) 114358.
66. Chen J, Yang Y, Song C, Zhang M, Wu S, Wang D (2019) Interfacial microstructure and mechanical properties of 316L/CuSn10 multi-material bimetallic structure fabricated by selective laser melting. Mater Sci Eng A 752:75–85
67. Pandiyan V, Baganis A, Richter RA, Wróbel R, Leinenbach C (2024) Qualify-as-you-go: sensor fusion of optical and acoustic signatures with contrastive deep learning for multi-material composition monitoring in laser powder bed fusion process. Virt Phys Prototyp 19:e2356080
68. Kavas B, Richter RA, Tucker MR, Pandiyan V (2025) Adaptive in-situ monitoring for laser powder bed fusion: self-supervised learning for layer thickness monitoring across scan lengths based on pyrometry. Opt Laser Technol 192:114070