Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 6, December 2026, pp. 5421~5437
ISSN: 2302-9285, DOI: 10.11591/eei.v15i6.13123

Hybrid transformer architecture for detecting anomalous z-offset regimes in FFF/FDM 3D printing

Arailym Kenzhebay[1], Ardak Karipzhanova[1], Nazira Ospanova[2], Alibek Barlybayev[3,4], Aigerim Bakiyeva[5], Anargul Bekenova[6], Yerbol Yerbayev[6], Assemgul Tynykulova[3]
[1]Faculty of Information Technology and Economics, Alikhan Bokeikhan University, Semey, Kazakhstan
[2]Department of Information Technologies, Non-profit Joint Stock Company «Toraighyrov University», Pavlodar, Kazakhstan
[3]Higher School of Information Technology and Engineering, Astana International University, Astana, Kazakhstan
[4]Faculty of Information Technology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
[5]Department of Software Engineering, Astana IT University, Astana, Kazakhstan
[6]Polytechnic Institute, Non-profit Joint-Stock Company West Kazakhstan Agrarian and Technical University named after Zhangir Khan, Uralsk, Kazakhstan

Article Info

Article history:

Received Apr 18, 2026
Revised Sep 21, 2026
Accepted Oct 4, 2026

Keywords:

Fused deposition modeling
Fused filament fabrication
Machine learning
Multichannel telemetry
Predictive quality control

ABSTRACT

This study proposes generalized adaptive telemetry encoder for 3D printing (GATE3D), a hybrid framework for early anomaly prediction in fused filament fabrication (FFF)/fused deposition modeling (FDM) additive manufacturing using only native printer telemetry. The framework combines temporal neural representations with gradient-boosted features and was tested on the Caxton dataset using a leak-free print group partition. On an unseen-print holdout comprising 5,323 windows from 27 previously unseen prints, GATE3D achieved Macro-F1 scores of 0.887, 0.883, and 0.905 for flow rate, feed rate, and hotend, respectively, while z-offset remained near chance level (0.503). Paired bootstrap analysis of print groups showed statistically confirmed improvement over HistGradientBoosting (HGB) for the hotend only. Event-level assessment further identified the hotend as the most reliable early warning target, with an event recall rate of 0.687 and 15.714 false positives per 1,000 stable nominal windows. Overall, GATE3D provides a competitive camera-free approach to telemetry-based early anomaly forecasting, while its effectiveness remains strongly target-dependent.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Nazira Ospanova
Department of Information Technologies, Non-profit Joint Stock Company «Toraighyrov University»
Pavlodar, Kazakhstan
Email: ospanova.n@teachers.tou.edu.kz

1. INTRODUCTION

Fused filament fabrication (FFF), also known as fused deposition modeling (FDM), has become one of the most widely used additive manufacturing technologies due to its low production costs, design flexibility, and the ability to custom-fabricate components with complex geometries. Despite these advantages, maintaining consistent print quality remains a significant challenge, particularly during the first layer deposition stage, where small variations in extrusion speed, nozzle movement, temperature conditions, or nozzle-to-build platform distance can propagate to subsequent layers, leading to dimensional inaccuracies, degraded surface quality, reduced mechanical strength, and even complete print failure [1]–[3]. Consequently, reliable online process monitoring and early prediction of abnormal operating conditions have become important research areas in the field of intelligent additive manufacturing.

Recent advances in artificial intelligence have significantly improved FFF/FDM printing process monitoring. Machine vision-based approaches remain the dominant research direction, with convolutional

neural networks and, more recently, vision transformer (ViT) architectures demonstrating remarkable results in detecting print defects, assessing layer quality, and recognizing extrusion anomalies directly from camera images [4]–[7]. Large-scale datasets such as Caxton, combined with Transformer-based visual recognition models, have significantly improved automatic quality assessment and defect localization [5], [6]. However, these methods require additional hardware for image processing, controlled lighting, camera calibration, and significant computational resources, increasing deployment costs and reducing scalability in industrial settings.

To overcome these limitations, recent studies have explored thermal imaging, acoustic sensing, vibration analysis, and multimodal sensor data fusion for monitoring additive manufacturing [8]–[10]. These approaches typically improve monitoring accuracy by integrating additional information sources, but require additional sensors and printer-specific calibration procedures, making practical implementation more complex and costly. Consequently, increasing attention has been paid to telemetry-based monitoring, where process variables generated directly by the printer—including extrusion speed, feed rate, nozzle and platform temperatures, and motion coordinates—are used to characterize process dynamics without relying on external sensor systems [11]–[13].

Recent transformer-based sequence learning models have demonstrated superior capabilities for modeling long-term temporal dependencies in industrial multivariate time series compared to traditional recurrent neural networks and classical machine learning algorithms [14]–[16]. These advances suggest that multichannel printer telemetry contains sufficient information to support predictive process monitoring. However, most existing telemetry-based studies focus on ongoing process state classification or the reconstruction of discretized process classes from their corresponding continuous variables. Such experimental conditions can lead to leakage of target information, as the target variable can be directly reconstructed from its original parameter rather than inferred from independent telemetry. Moreover, a number of studies claiming camera-free monitoring still use statistical descriptors derived from visual observation images, which cannot be considered strictly non-visual telemetry. These methodological issues reduce the practical applicability and reproducibility of the reported results and complicate the quantitative comparison of different approaches [17]–[19].

Another important limitation concerns the formulation of the prediction problem. Previous studies often use the terms "process state recognition," "parameter classification," "anomaly detection," and "early failure prediction" interchangeably, although they represent fundamentally different learning tasks. In practical applications, predictive monitoring should estimate future anomalous process conditions using only information available at the time of inference, rather than reconstructing the current state based on directly related variables. Accordingly, this study proposes generalized adaptive telemetry encoder for 3D printing (GATE3D), a hybrid framework for early anomaly prediction based on a strict leakage-free protocol that excludes target-related variables and all image-derived features. The proposed architecture combines time series learning with a nonlinear telemetry representation and is evaluated against classical machine learning models, ensemble, recurrent, and transformer models using identical pre-processing, group data splitting, and standardized evaluation metrics.

Unlike existing studies on Transformer-based FFF/FDM monitoring, which mainly focus on visual defect recognition, current state classification, or parameter reconstruction, the novelty of GATE3D lies in formulating telemetry-only monitoring as a leakage-controlled early anomaly prediction problem, in which future abnormal states are inferred solely from information available up to the prediction horizon. The main contributions of this study are as follows. First, a leakage-controlled formulation for early anomaly prediction based on multi-channel FFF/FDM telemetry is presented, in which future abnormal operating conditions are predicted solely based on information available up to the forecast horizon, while variables directly related to the prediction targets are excluded. Second, a monitoring protocol is developed that completely eliminates the use of cameras by removing all image-derived features and retaining only the printer's internal telemetry, providing a less complex alternative to machine vision-based and multimodal monitoring systems. Third, GATE3D presents a hybrid prediction architecture that integrates learning a temporal representation of multi-channel telemetry with nonlinear modeling of the interaction of process variables, enabling the system to account for both time-dependent process dynamics and complex interrelations between telemetry signals within a single prediction pipeline. Fourth, a reproducible benchmarking protocol was developed to evaluate GATE3D against classical machine learning models, ensemble, recurrent, convolutional, and transformer models under identical preprocessing, print-level data partitioning, and evaluation conditions. Taken together, these factors distinguish the proposed approach from traditional methods of current state or parameter recovery and provide a practical basis for predicting early anomalies in FFF/FDM processes with leakage control and using telemetry alone.

The remainder of the paper is organized as follows. Section 2 introduces the problem statement, the proposed GATE3D framework, and the experimental protocol, including a leak-free data preparation strategy and implementation parameters. Section 3 presents a preliminary analysis of telemetry data and a comparative performance evaluation of GATE3D against the baseline and state-of-the-art models. Section 4 discusses the

experimental results, highlights the advantages and limitations of the proposed approach, and outlines directions for future research. Finally, section 5 concludes the paper, summarizing the main achievements and practical implications of the proposed framework.

2. METHOD

Existing approaches to FFF/FDM monitoring can be roughly divided into four categories: visual, thermal/infrared, acoustic or multimodal methods, and direct parameter classification. Visual methods provide effective defect detection but require cameras, controlled lighting, and calibration. Thermal and multimodal approaches improve process observability but rely on additional sensors and specialized equipment. Direct parameter classification predicts the current process state based on relevant telemetry; however, it can lead to target leakage when reconstructing discretized classes from their original continuous variables. The main symbols and variables used in the mathematical formulation of the proposed method are given in Table 1.

Table 1. Notation used in the proposed method

Symbol	Definition
p	Identifier of an individual print (print_id)
t	Current frame/time index
L	Length of the input temporal window
H	Forecasting horizon
q	Technological prediction target
$x_{p,t}$	Vector of permitted input features for print (p) at time (t)
D	Number of sequential input feature channels
$X_{p,t}$	Input window consisting of (L) consecutive feature vectors
$c_{p,t+H}^{(q)}$	Original technological class of target (q) at forecasting horizon (H)
$y_{p,t}^{(q)}$	Binary future label for target (q): 0 denotes a future nominal state and 1 denotes a future non-nominal state
$h_{p,t}$	Hidden temporal representation produced by the sequential encoder
α_t	Normalized attention-pooling weight
$s_{p,t}$	Aggregated summary representation of the input window
$p_{\mathrm{N}}^{(q)}$	Predicted probability produced by the neural branch for target (q)
$p_{\mathrm{B}}^{(q)}$	Predicted probability produced by the gradient-boosted branch for target (q)
λ_q	Fusion coefficient for target (k), selected using the validation set only
τ_q	Classification threshold for target (q), selected using the validation set only

In this paper, a different configuration is considered. The model receives only a time window of current and past non-visual telemetry and predicts whether an anomalous regime will occur on the positive time horizon. Let p denote the seal identifier, t be the current time index, L=69 be the window length, H=30 be the forecast horizon, and k be a technological target from the set (1):

$$K = \{\text{flow_rate, feed_rate, z_offset, hotend}\} \quad (1)$$

The input sequence is defined as (2):

$$X_{p,t} = [x_{p,t-L+1}, x_{p,t-L+2}, \dots, x_{p,t}] \quad (2)$$

where $X_{p,t}$ contains only valid non-visual features available before the prediction time. The binary future target variable is defined as (3):

$$y_{p,t+H}^{(k)} = \begin{cases} 0, & c_{p,t+H}^{(k)} = 1, \\ 1, & c_{p,t+H}^{(k)} \in \{0,2\} \end{cases} \quad (3)$$

where $c_{p,t+H}^{(k)}$ is the future technological class of the corresponding parameter. Thus, class 0 denotes the future nominal mode, and class 1 denotes the future transition to an abnormal mode. This formulation excludes the restoration of the current class from its original variable and corresponds to the problem of early forecasting of the future technological state.

The proposed GATE3D model is based on the assumption that the future anomalous mode of FFF/FDM printing is determined by two complementary types of information: the short-term temporal dynamics of telemetry and the nonlinear state of the process mode within the window. Therefore, GATE3D includes a temporal neural branch and a boosted regime branch. The temporal branch transforms the input sequence $X_{p,t}$ into a latent temporal representation (4):

$$Z = \phi_\theta(X_{p,t}) \tag{4}$$

where ϕ_θ includes normalization, linear projection, one-dimensional temporal convolution blocks [20], residual links, and attention pooling. Attention pooling is defined as (5) and (6):

$$\alpha_i = \frac{\exp\left(v^\top \tanh\left(W z_i\right)\right)}{\sum_{j=1}^{L} \exp\left(v^\top \tanh\left(W z_j\right)\right)} \tag{5}$$

$$h_{temp} = \left[\sum_{i=1}^{L} \alpha_i \, ; z_i ; z_L \, \frac{1}{L} \sum_{i=1}^{L} z_i\right] \tag{6}$$

This representation combines three components: attention-weighted state, the latest window state, and the average window dynamics. This allows for consideration of both local changes and the overall process behavior. The boosted regime branch constructs an aggregated description of the window (7):

$$s_{p,t} = g(X_{p,t}) \tag{7}$$

where $g(\cdot)$ includes window statistics, lagged features, rolling mean, rolling standard deviation, and trend features. For each target, the gradient-boosted branch [21], [22] estimates the probability of a future anomaly (8):

$$\hat{p}_{reg}^{(k)} = B_k(s_{p,t}) \tag{8}$$

The final GATE3D probability is formed through validation-based fusion (9):

$$\hat{p}^{(k)} = \lambda_k \hat{p}_{temp}^{(k)} + (1 - \lambda_k)\hat{p}_{reg}^{(k)} \tag{9}$$

where λ_k is selected only for the validation split and then fixed for the test and unseen-print holdout dataset. The loss function for the neural branch is defined as a weighted multi-task binary cross-entropy (10):

$$\mathcal{L} = \frac{1}{N|\mathcal{K}|} \sum_{i=1}^{N} \sum_{k \in \mathcal{K}} \omega_k \left[-\beta_k y_i^{(k)} \log \hat{p}_i^{(k)} - \left(1 - y_i^{(k)}\right) \log\left(1 - \hat{p}_i^{(k)}\right)\right] \tag{10}$$

where β_k compensates for class imbalance and ω_k specifies the weight of the technological goal.

The theoretical advantage of GATE3D is that the temporal branch better describes the order and dynamics of events, while the boosted regime branch reliably extracts nonlinear tabular dependencies. Combining them reduces the risk of a one-sided model dependence on sequence patterns alone or on aggregated features alone. The study used the CAXTON dataset (Table 2), which contains 948,396 sequential telemetry observations. Future labels were formed separately within each print_id over a fixed horizon of H=30 frames. 5,441 boundary observations were deterministically excluded solely because there was no state for them at position t+30 within the corresponding print. After this operation, 942,955 model observations remained. No additional row deletion was performed when reading data or parsing timestamps.

Table 2. Data preparation, group-aware partitioning, and window-generation protocol

Component	Values
Raw observations	948,396
Excluded at future-label construction	5,441
Modeling observations	942,955
Forecast horizon	30 frames
Sequence length	69 frames
Base telemetry channels	7
Sequence input channels	38
Summary representation	798 train-fitted features
Train	564,111 rows; 114 print_id; and 22,000 windows
Validation	135,824 rows; 19 print_id; and 4,812 windows
Test	92,486 rows; 20 print_id; and 3,266 windows
Unseen-print holdout	150,534 rows; 27 print_id; and 5,323 windows
Train stride	16
Validation/test/holdout stride	28
Train overlap	76.8116%
Evaluation overlap	59.4203%
Primary metric	Macro-F1
Threshold selection	Validation only

The main sequential models use a window of length L=69 frames. The basic non-visual set contains seven process channels: flow_rate, feed_rate, target_hotend, hotend, bed, nozzle_tip_x, and nozzle_tip_y. Past-only lag, rolling mean, rolling standard deviation, and trend features are calculated for them. Together with the temporal context t_rel_sec, layer_idx, and frame index, the sequential representation contains 38 channels. The variables z_offset, current process mode classes, future labels, img_mean, and img_std are excluded from the predicted inputs. img_num is used only as a frame index and is referred to as frame_index in the manuscript.

Table 3 shows the complete window generation procedure for each data partition. For reference, the table lists the theoretical number of windows obtained with a stride of 1, the number of windows remaining after applying a predefined sampling step, and the final number of windows used for training and evaluating the model.

Table 3. Audit of time window formation

Split	Rows	Windows at step 1	After a given step	After the restriction	Step	Overlap (%)
Train	564,111	556,359	34,824	22,000	16	76.81
Validation	135,824	134,532	4,812	4,812	28	59.42
Test	92,486	91,126	3,266	3,266	28	59.42
Unseen-print holdout	150,534	148,698	5,323	5,323	28	59.42

Window generation was performed independently within each print_id, preventing windows from crossing print boundaries. A stride of 16 was used for the training set, while a stride of 28 was used for the validation, test, and control sets with invisible fingerprints. To reduce computational cost while maintaining representative temporal variability, the training set was further limited to 22,000 windows, while all windows in the evaluation partitions were retained after stride-based sampling. The corresponding overlap coefficients are also provided, providing a complete and reproducible description of the temporal sampling strategy.

3. RESULTS

This section presents an analysis of FFF/FDM telemetry data and an experimental evaluation of the proposed GATE3D framework. First, the distributions of the input and target variables are examined, followed by the linear and nonlinear feature dependencies to characterize the dataset and evaluate its predictive potential. Finally, GATE3D is compared to traditional machine learning and deep learning models within a single experimental protocol. Figure 1 shows the empirical class distribution of the flow_rate_class variable, which describes the material flow rate during 3D printing. It is evident that the largest share of observations is made up of states with reduced (class 0, about 341,000 points) and increased supply (class 2, about 371,000 points), while the nominal mode (class 1, about 236,000 points) occurs significantly less frequently. Before comparing the forecasting models, the experimental results first test the practical implications of the strict leakage-free protocol presented in section 2. The comparison includes classical machine learning methods, boosting algorithms, recurrent neural networks [23], temporal convolutional networks [20], and transformer-based architectures [14], evaluated under exactly the same preprocessing conditions, forecasting horizon, data partitioning strategy, and evaluation metrics.

This ratio suggests that the initial experiments intentionally exceeded the optimal extrusion settings to explore extreme process conditions. For subsequent model training, this distribution indicates a moderate class imbalance and the need to control for the influence of mode frequency when selecting quality metrics and validation strategies. To assess the structure of the target variables, class distributions for each process parameter were analyzed. As shown in Figure 1(a), the flow_rate_class variable has a relatively balanced distribution between classes 0 and 2, while class 1 is represented by a smaller number of observations. A similar trend is observed for feed_rate_class (Figure 1(b)), where classes 0 and 2 contain a comparable number of records, and class 1 is less common. The most pronounced imbalance was found for z_offset_class (Figure 1(c)), in which class 2 significantly predominates over classes 0 and 1. In turn, for hotend_class (Figure 1(d)), the largest number of observations belong to class 1, while classes 0 and 2 are represented by significantly fewer instances. The resulting distributions indicate the presence of varying degrees of class imbalance, which should be taken into account when training and evaluating predictive models.

For a more detailed analysis of the source data, the distributions of the continuous process parameters used to generate the target classes were examined, as shown in Figure 2. Figure 2(a) shows the distribution of the flow_rate parameter, which characterizes the material feed rate. Values are concentrated primarily in the range from 0 to 200 units, with an average value of 111 and a median of 111, indicating a virtually symmetrical distribution. The distribution of the feed_rate parameter (Figure 2(b)) covers the range from 20 to 200 units and is characterized by uniform coverage of the value interval with an average value of 110 and a median of 110. The z_offset parameter (Figure 2(c)), which reflects the deviation in the distance between the

nozzle and the print surface, is distributed in the range from -0.08 to 0.32 mm with an average value of 0.119 and a median of 0.120, indicating the absence of significant data bias. For the target_hotend extruder temperature (Figure 2(d)), a concentration of values in the range of 180–230 °C is observed, with a distinct peak at approximately 206 °C, an average of 207 °C, and a median of 206 °C. The resulting distributions demonstrate a sufficient diversity of printing operating modes and confirm the representativeness of the telemetry data for subsequent modeling and process prediction.

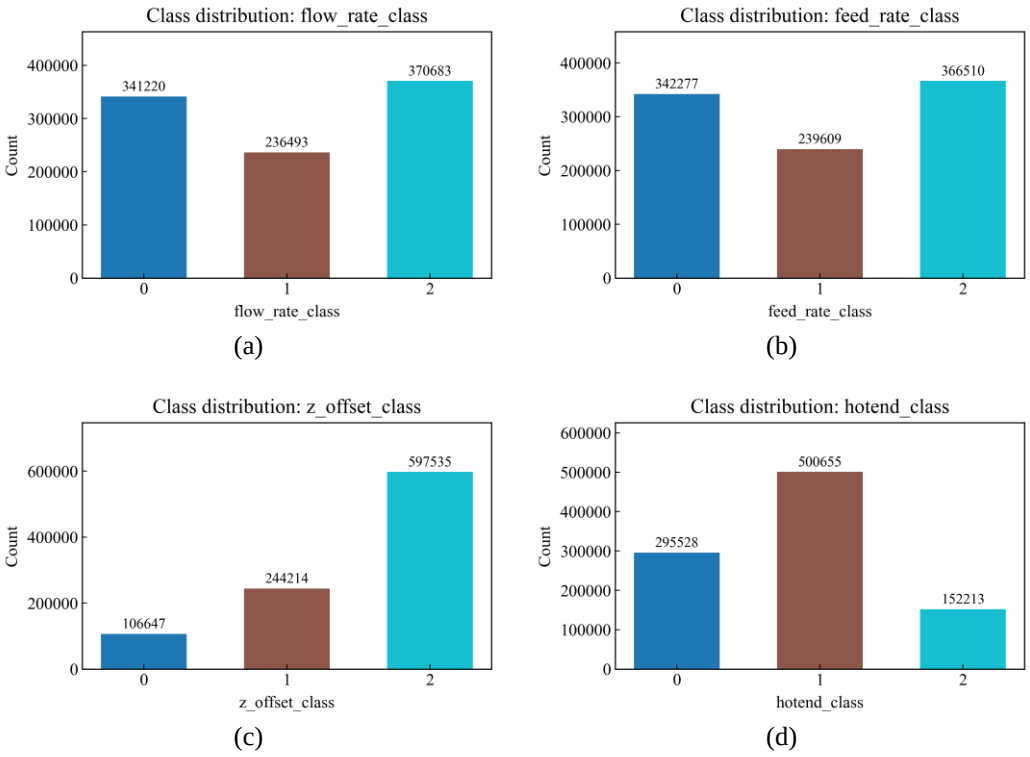


Figure 1. Distribution of objective function classes; (a) flow_rate_class, (b) feed_rate_class, (c) z_offset_class, and (d) hotend_class

Additionally, the distributions of the print head temperature and spatial position parameters, which directly impact the stability of the additive manufacturing process, were analyzed, as shown in Figure 3. Figure 3(a) shows the distribution of the actual hotend nozzle temperature, which exhibits a nearly normal distribution shape with a concentration of values in the 180–230 °C range. The mean value is 204 °C, and the median is 204 °C, indicating the absence of a noticeable shift. The bed temperature parameter (Figure 3(b)) is characterized by a pronounced maximum of approximately 65 °C with an average value of 64.8 °C and a median of 65.1 °C, reflecting the predominance of printing modes with constant heating of the working surface. The spatial coordinates of the nozzle position along the X axis (nozzle_tip_x, Figure 3(c)) and the Y axis (nozzle_tip_y, Figure 3(d)) demonstrate multimodal distributions due to the peculiarities of the print head movement trajectory during the formation of the product. For nozzle_tip_x, the mean is 554 and the median is 570, while for nozzle_tip_y, the mean is 530 and the median is 533. The presence of several local maxima indicates uneven use of the printer's working area and confirms the informativeness of coordinate features for analyzing the state of the printing process and predicting possible deviations.

Figure 4 shows the Pearson correlation matrix for the initial non-visual process variables and process class labels. This analysis was conducted only for preliminary characterization of the Caxton dataset and was not used as part of the final set of predictive features. In the final protocol, without the use of cameras and data leaks, image-derived variables such as img_mean and img_std were excluded, and the predictive models used only the printer's own telemetry data available at the time of prediction.

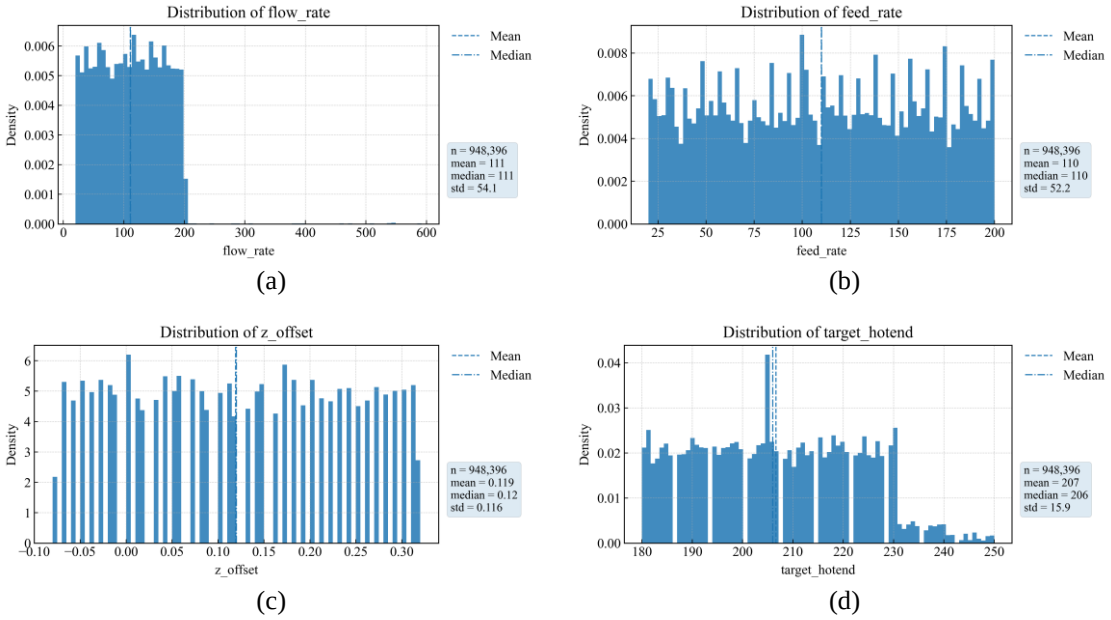


Figure 2. Distributions of continuous process parameters; (a) flow_rate, (b) feed_rate, (c) z_offset, and (d) target_hotend

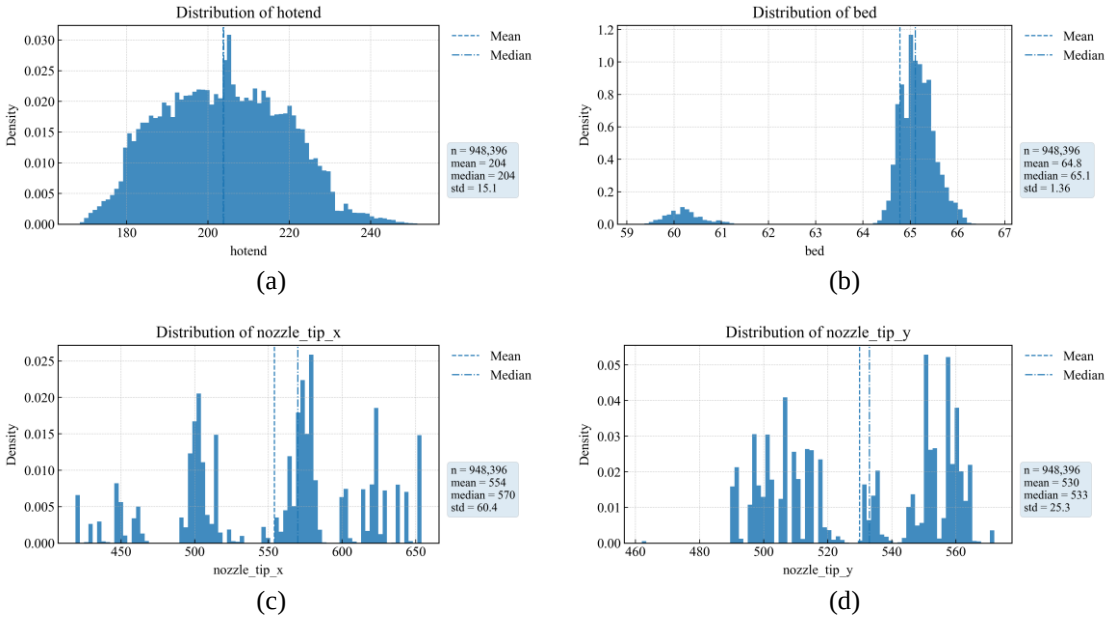


Figure 3. Distributions of temperature and spatial parameters of the printing process; (a) hotend, (b) bed, (c) nozzle_tip_x, and (d) nozzle_tip_y

The features×classes block shows that the continuous variables flow_rate, feed_rate, z_offset, and hotend are highly correlated with their respective target classes (flow_rate_class, feed_rate_class, z_offset_class, and hotend_class), with Pearson coefficients ranging from 0.76 to 0.94. This confirms that each target class represents a discretized version of its corresponding process variable, while cross-correlations between unrelated variables remain insignificant. Figure 5 presents the results of an XGBoost-based feature importance analysis, where each variable was predicted using all remaining features and class labels. The results show that the discretized target classes are almost completely recovered from their corresponding continuous variables (importance=0.90–1.00), although the opposite relationship is also evident, with continuous variables being mostly explained by their own classes. These results confirm that the target labels are threshold-based transformations of the original process variables and, therefore, can lead to target information leakage if both representations are used simultaneously.

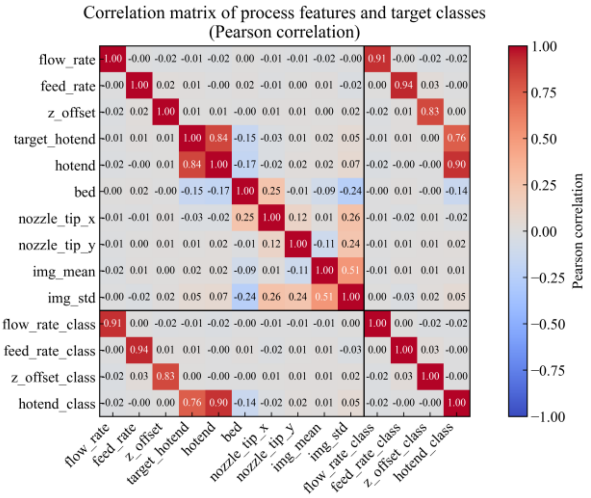


Figure 4. Pairwise correlation matrix of technological features and target classes of the 3D printing process (Pearson coefficient)

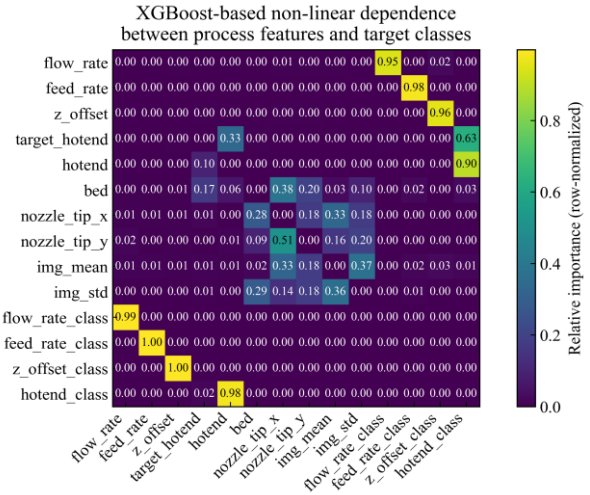


Figure 5. Matrix of nonlinear interdependence of technological features and target classes based on XGBoost estimates

Within the block of continuous features, a distinct group related to bed temperature and spatial/diagnostic parameters stands out. For the "bed" row, the contribution of nozzle coordinates (nozzle_tip_x, nozzle_tip_y) and the features img_mean and img_std is noticeable, while for the latter, both bed and coordinates are significant. This pattern of importance indicates a nonlinear relationship between the position of the print zone, thermal conditions on the bed, and the nature of the diagnostic signal. For the remaining combinations, the significance is close to zero, indicating the absence of pronounced nonlinear dependencies between most telemetry channels and "foreign" target classes. Taken together, the matrix emphasizes the dual nature of the data structure: on the one hand, the classes directly encode their own continuous parameters, while on the other, interchannel interactions are localized in a narrow subset of features related to the spatial-thermal state of the build zone. Figure 6 shows the distance correlation matrix (dCorr) [24] for a set of continuous process parameters and their corresponding target classes.

Unlike the classic Pearson coefficient, the dCorr metric is sensitive to any type of statistical dependence, including nonlinear and nonmonotonic ones, allowing for a more complete description of the relationships between telemetry channels. Diagonal elements equal to one correspond to a trivial relationship between the variable and itself. In the "feature–eigenclass" block, very high dCorr values are recorded: the pairs flow_rate–flow_rate_class, feed_rate–feed_rate_class, and z_offset–z_offset_class demonstrate

coefficients of approximately 0.87-0.95, while the combinations target_hotend–hotend_class and hotend–hotend_class reach 0.73 and 0.88, respectively. This confirms that the discrete mode labels are closely, but not strictly linearly, related to the original continuous parameters and effectively encode their ranges. In the block of relationships between the features themselves, groups of parameters participating in joint dynamics are clearly visible. The temperature channel is characterized by a high dCorr between target_hotend and hotend (~0.82), reflecting the coordinated operation of the control loop. The spatial diagnostic cluster is defined by the pairs nozzle_tip_x–nozzle_tip_y (~0.45), nozzle_tip_x–img_mean and nozzle_tip_y–img_mean (0.39–0.40), and img_mean–img_std (~0.49). Non-zero coefficients between bed and these same features (0.20–0.23) indicate a more complex than linear relationship between the quality of the diagnostic signal and the thermal state of the bed and the print zone position.

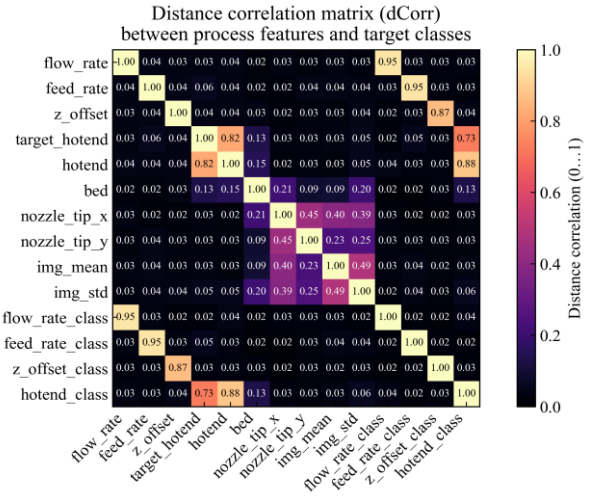


Figure 6. The matrix of dCorr dependencies between technological features and target classes of the 3D printing process

For most other combinations, dCorr values are in the range of 0.02–0.06, indicating a weak or absent stable relationship between the corresponding channels. Overall, the matrix demonstrates that, in addition to the practically deterministic parameter-class relationships, the data contains compact groups of variables with pronounced nonlinear interdependencies, which can potentially be used by artificial intelligence models to more accurately predict abnormal process behavior. Figure 7 shows the mutual information matrix, which estimates the amount of shared information shared by pairs of process features and target classes. The values are presented in NATs and allow us to assess the strength of both linear and nonlinear relationships between variables. In the feature-feature block, a spatial diagnostic cluster is clearly visible. The nozzle_tip_x–nozzle_tip_y coordinate pair exhibits the maximum mutual information value (approximately 3.26 nata), reflecting the tight coupling of the X- and Y-axis motion trajectories. Comparable, albeit slightly lower, values are observed for the nozzle_tip_x–img_mean, nozzle_tip_x–img_std, as well as nozzle_tip_y–img_mean and nozzle_tip_y–img_std pairs (0.53–0.72 nata). This indicates pronounced spatial heterogeneity of the diagnostic signal: the characteristics of the observed print zone depend significantly on the nozzle position on the table. Among the process parameters, the flow_rate–feed_rate (~1.48 nata) and target_hotend–hotend (~1.31 nata) relationships are noticeable, demonstrating consistent material feed adjustment and the efficient operation of the nozzle temperature circuit. The "feature-class" block shows that discrete target variables carry a significant amount of information about the corresponding continuous parameters. For example, the flow_rate–flow_rate_class and feed_rate–feed_rate_class pairs have a mutual information of approximately 1.08 nat, z_offset–z_offset_class has a mutual information of approximately 0.88 nat, and hotend–hotend_class has a mutual information of almost 1 nat. Thus, the mode classes effectively encode an interval partitioning of the original quantities, while remaining sensitive to the nonlinear structure of their distributions. For the remaining combinations of features and classes, the mutual information values are close to zero, indicating their relative independence.

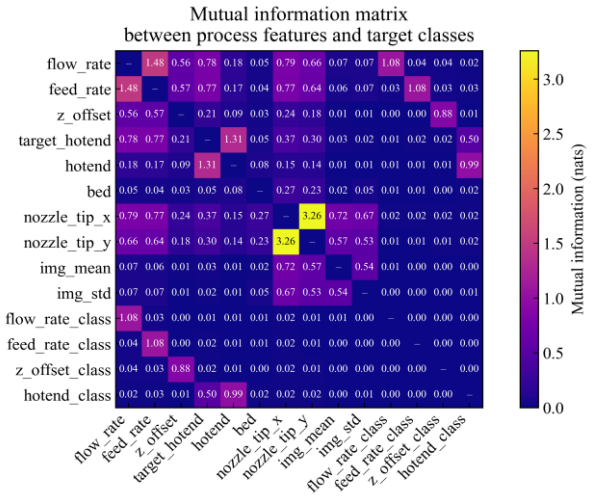


Figure 7. Matrix of mutual information between technological features and target classes of the 3D printing process

Table 4 compares GATE3D with traditional machine learning and sequence-based models on the same set of 5,323 windows containing no previously seen fingerprints. GATE3D achieved the highest Macro-F1 score among the strict prediction models for flow_rate (0.887), feed_rate (0.883), and hotend (0.905), outperforming HistGradientBoosting (HGB) (0.882, 0.881, and 0.902, respectively). For z_offset, all strict prediction models remained close to chance performance, with GATE3D achieving 0.503. Diagnostic persistence achieved 0.884 for z_offset, but serves only as a diagnostic control and not as a usable prediction model.

Table 4. Head-to-head Macro-F1 comparison on the unseen-print holdout

Model	Flow rate	Feed rate	Z-offset	Hotend
GATE3D	0.886603	0.882994	0.503279	0.905020
Diagnostic persistence*	0.884290	0.882036	0.883970	0.772689
Strong HGB	0.882411	0.881267	0.498401	0.901540
Strong ExtraTrees	0.870628	0.873140	0.505747	0.900899
TCN	0.852117	0.848121	0.473091	0.883191
LSTM	0.830168	0.827812	0.496407	0.873958
BiLSTM	0.825274	0.814420	0.499251	0.863659
PatchTST-style	0.824731	0.822473	0.500036	0.874289
GRU	0.821452	0.816674	0.504404	0.878374
AAT supervised H=30 adaptation	0.554475	0.542127	0.497652	0.771020
TransformerEncoder	0.505241	0.498837	0.506982	0.733448

Table 5 presents the final GATE3D performance on the unseen-print holdout comprising 5,323 windows from 27 print_ids that were not used during model training. All reported metrics were computed from a single frozen prediction release. The decision thresholds of 0.354, 0.589, 0.470, and 0.431 for flow_rate, feed_rate, z_offset, and hotend, respectively, were selected exclusively on the validation split and subsequently applied unchanged to the unseen-print holdout. GATE3D achieved Macro-F1 scores of 0.886603, 0.882994, and 0.905020 for flow_rate, feed_rate, and hotend, respectively. In contrast, z_offset achieved a Macro-F1 of 0.503279 and a balanced accuracy of 0.503820, indicating approximately chance-level discrimination.

Table 5. Authoritative GATE3D performance on the unseen-print holdout

Target	Threshold	Accuracy	Balanced accuracy	Macro-F1	PR-AUC	ROC-AUC	TN	FP	FN	TP
flow_rate	0.354	0.915649	0.883793	**0.886603**	0.981885	0.952686	1,090	239	210	3,784
feed_rate	0.589	0.913770	0.882694	**0.882994**	0.980310	0.949528	1,067	231	228	3,797
z_offset	0.470	0.605861	0.503820	**0.503279**	0.740105	0.492560	403	980	1,118	2,822
hotend	0.431	0.905129	0.905069	**0.905020**	0.965912	0.966069	2,499	257	248	2,319

Table 6 presents a comparison of paired bootstrap test [25] results for the printing groups between GATE3D and HGB based on 2,000 replicate samples. For flow rate and feed rate, GATE3D yielded slightly higher Macro-F1 values, with differences of +0.004193 and +0.001727, respectively; however, both 95% confidence intervals included zero, indicating no statistically supported difference. For z-bias, the difference was also nonsignificant (+0.004878; 95% CI: -0.018001 to 0.027208), with both models maintaining performance close to chance. Only for the hotend was the confidence interval completely above zero: GATE3D achieved 0.905020 compared to 0.901540 for HGB, with Δ=+0.003480 and 95% CI [0.000550; 0.006878]. Thus, a statistically confirmed advantage of GATE3D over HGB was observed only for the hotend.

Table 6. Paired print-group bootstrap

Target	GATE3D	HGB	Point Δ	Bootstrap mean Δ	95% CI
Flow	0.886603	0.882411	+0.004193	+0.003908	[−0.003013; 0.009667]
Feed	0.882994	0.881267	+0.001727	+0.001635	[−0.002149; 0.005513]
Z-offset	0.503279	0.498401	+0.004878	+0.005503	[−0.018001; 0.027208]
Hotend	0.905020	0.901540	+0.003480	+0.003570	[0.000550; 0.006878]

Table 7 summarizes the event-level performance across the validation runpoints. For flow_rate, the core threshold (0.354) detected 25 of 221 events (0.113 recall), increasing to 106 of 221 (0.480) at the 10% false-positive rate (FAR) runpoint. For feed_rate, event recall increased from 16 of 225 (0.071) to 113 of 225 (0.502) over the same range of runpoints. Hotend demonstrated the best event-level performance, detecting 480 of 699 events (0.687 recall) at the core threshold and maintaining a recall of 0.648 (453 of 699) at approximately 1% validation FAR. In contrast, z_offset demonstrated an apparent recall of 0.692 (157/227) at its primary threshold, but with a very high FAR in the control group of 0.713 (712.794 false positives per 1,000 stable nominal windows). At a performance FAR of 1%, z_offset detected only 1/227 events, confirming weak event-level discrimination under false-positive-controlled conditions.

Table 7. Event operating points selected exclusively on validation

Target	Operating point	Threshold	Validation FAR	Holdout FAR	Detected/events	Event recall	False alerts/1,000 stable nominal
flow	Macro-F1	0.354000	0.007701	0.010018	25/221	0.113122	10.018
flow	FAR 1%	0.351489	0.009901	0.011840	26/221	0.117647	11.840
flow	FAR 2.5%	0.324147	0.024202	0.020036	35/221	0.158371	20.036
flow	FAR 5%	0.287564	0.049505	0.045537	55/221	0.248869	45.537
flow	FAR 10%	0.233386	0.099010	0.089253	106/221	0.479638	89.253
feed	Macro-F1	0.589000	0.004028	0.001878	16/225	0.071111	1.878
feed	FAR 1%	0.570295	0.009063	0.002817	19/225	0.084444	2.817
feed	FAR 2.5%	0.503161	0.024169	0.025352	39/225	0.173333	25.352
feed	FAR 5%	0.471920	0.049345	0.057277	60/225	0.266667	57.277
feed	FAR 10%	0.419204	0.099698	0.092958	113/225	0.502222	92.958
z_offset	Macro-F1	0.470000	0.716216	0.712794	157/227	0.691630	712.794
z_offset	FAR 1%	0.868755	0.009653	0.001741	1/227	0.004405	1.741
z_offset	FAR 2.5%	0.833222	0.024131	0.018277	2/227	0.008811	18.277
z_offset	FAR 5%	0.797398	0.049228	0.044386	3/227	0.013216	44.386
z_offset	FAR 10%	0.756374	0.099421	0.090513	5/227	0.022026	90.513
hotend	Macro-F1	0.431000	0.011117	0.015714	480/699	0.686695	15.714
hotend	FAR 1%	0.532610	0.009529	0.011905	453/699	0.648069	11.905
hotend	FAR 2.5%	0.314100	0.024881	0.020952	512/699	0.732475	20.952
hotend	FAR 5%	0.221480	0.049762	0.043810	553/699	0.791130	43.810
hotend	FAR 10%	0.171295	0.099524	0.081905	581/699	0.831187	81.905

Table 8 presents the results of the GATE3D branch analysis. The full model achieved a higher Macro-F1 score than each individual branch for flow (0.887), feed (0.883), and hotend (0.905), with gains over the boosted-only branch of +0.002, +0.002, and +0.005, respectively. The largest improvement among successfully predicted targets was observed for the hotend, indicating the most pronounced contribution of hybrid fusion. For z_offset, the full model achieved 0.503, compared to 0.490 and 0.486 for the neural and boosted branches; however, all configurations remained close to random performance. Overall, the analysis suggests that the contribution of hybrid fusion is target-dependent and not uniformly beneficial across all targets.

Table 8. Primary branch-level ablation on the unseen-print holdout

Variant	Flow	Feed	Z-offset	Hotend
Neural branch only	0.833540	0.818900	0.490254	0.872055
Boosted branch only	0.884654	0.880841	0.485668	0.900009
Full GATE3D	0.886603	0.882994	0.503279	0.905020
Full−boosted	+0.001949	+0.002153	+0.017611*	+0.005011

Table 9 shows the key event-level performance metrics of GATE3D using the invisible fingerprint control group. The hotend demonstrated the best event detection performance, identifying 480 of 699 events with a recall rate of 0.687 [95% CI: 0.624–0.744], an event detection accuracy of 0.186, and 15.714 false positives per 1,000 stable nominal windows. Flow and feed achieved lower event detection recall rates: 0.113 (25/221) and 0.071 (16/225), respectively. Although z_offset demonstrated a high event detection recall rate of 0.692 (157/227), this was accompanied by 712.794 false positives per 1,000 windows, indicating an impractical operating point. The median lead time ranged from 6.44 s for the hotend to 7.82 s for the stream and z_offset. Overall, the hotend system provided the most favorable balance between event detection and false positive rates.

Table 9. Principal event-level operating characteristics on the unseen-print holdout

Target	Events	Detected	Event recall (95% CI)	One-to-one event precision (95% CI)	Alarms/print (95% CI)	False alerts/1,000	Median lead
Flow	221	25	0.1131 [0.0758; 0.1560]	0.0062 [0.0038; 0.0090]	149.0 [84.36; 223.85]	10.018	7.82 s
Feed	225	16	0.0711 [0.0379; 0.1067]	0.0040 [0.0020; 0.0073]	149.19 [82.70; 224.74]	1.878	7.59 s
Z-offset	227	157	0.6916 [0.6384; 0.7530]	0.0413 [0.0362; 0.0467]	140.81 [79.66; 207.97]	712.794	7.82 s
Hotend	699	480	0.6867 [0.6238; 0.7442]	0.1863 [0.1668; 0.2067]	95.41 [54.63; 141.08]	15.714	6.44 s

Figure 8 shows the event precision and recall curves as a function of time for flow velocity, feed rate, Z-axis offset, and hotend at lead times of 1, 5, 10, 15, 20, and 30 frames. The maximum event precision was 0.046 for flow velocity, 0.047 for feed rate, 0.043 for Z-axis offset, and 0.197 for hotend. Increasing the required lead time generally reduced both the achievable event recall and precision, especially at 20–30 frames. The hotend demonstrated the highest and most stable event-level performance, while flow velocity, feed rate, and Z-axis offset showed relatively low event precision. These results illustrate the tradeoff between early warning lead times and reliable event detection for the four prediction targets.

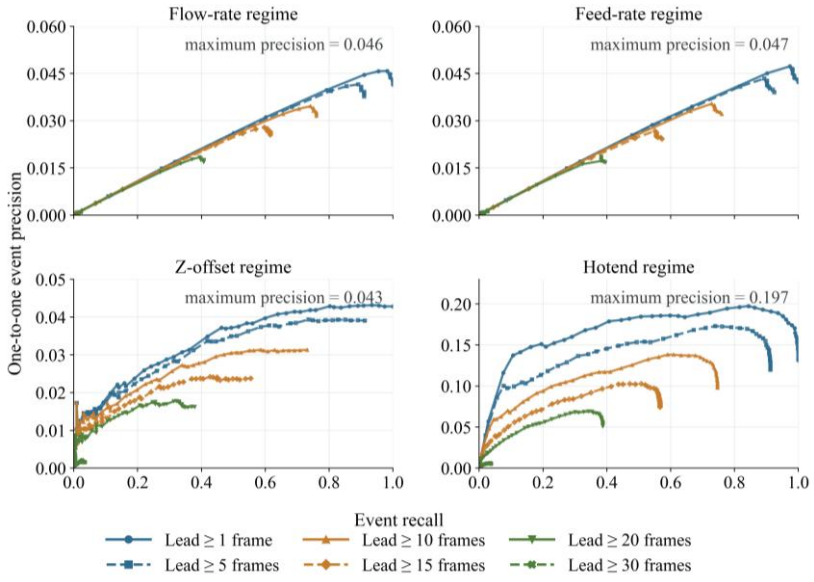


Figure 8. Time curves of precision–recall of events at all thresholds

Figure 9 presents the sensitivity of GATE3D and HGB to the asymmetric costs of false negative and false positive errors at FN:FP ratios of 1, 2, 5, 10, and 20. The thresholds were selected exclusively on the validation set and subsequently evaluated on the unseen-print holdout. HGB showed lower relative costs for flow rate at ratios of 2–20 and for z-axis offset at all ratios evaluated. For feed rate, both models showed generally comparable costs, while HGB had an advantage at a ratio of 5. For the hotend, GATE3D showed an advantage at a ratio of 1, while the differences at higher ratios were insignificant.

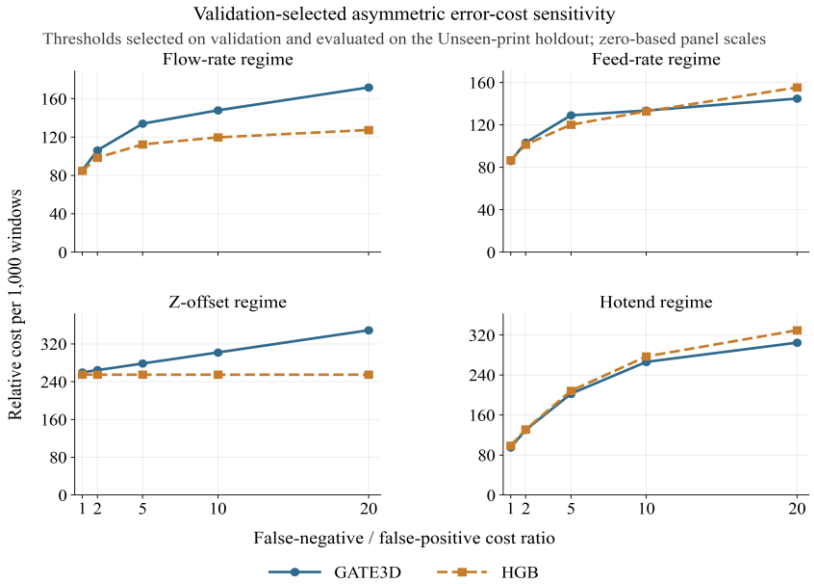


Figure 9. Sensitivity to asymmetric error costs at validation-selected thresholds

Overall, the experimental results showed that GATE3D demonstrated high window-level prediction accuracy for flow rate, feed rate, and heating element, with Macro-F1 scores of 0.8866, 0.8830, and 0.9050, respectively, while the z-bias remained close to chance (0.5033). Paired bootstrap analysis for print groups confirmed a statistically significant advantage over HGB only for the heating element (Δ=+0.0035; 95% CI: 0.0006–0.0069), while the differences for flow rate and feed rate were insignificant. Event-level evaluation further identified the heater as the most effective early warning target, detecting 480 of 699 events with a recall of 0.6867, while the apparent high event recall for z-shift was accompanied by 712.794 false positives per 1,000 stable nominal windows. Taken together, these results demonstrate significant differences in window-level prediction and event detection performance depending on the target.

4. DISCUSSION

The results demonstrate that integrated non-visual printer telemetry can support early prediction of individual FFF/FDM process anomalies within a strict, leak-free protocol. On a control set of invisible prints, GATE3D achieved Macro-F1 scores of 0.8866, 0.8830, and 0.9050 for flow rate, feed rate, and heating element, respectively, while the z-axis offset remained close to chance (0.5033). GATE3D achieved the highest Macro-F1 score among the strict prediction models for flow rate, feed rate, and heating element. However, a paired bootstrap analysis of the print groups revealed that its advantage over HGB was statistically confirmed only for the heating element (Δ=+0.0035; 95% CI: 0.0006–0.0069). The differences for flow rate and feed rate were not statistically significant. Similarly, ablation analysis revealed that hybrid fusion provided the most significant improvement for the hotend (+0.0050 compared to the boosted branch), indicating that the hybrid architecture's contribution is dependent on target performance.

Event-level evaluation further emphasized this target-dependent performance. Hotend demonstrated the best early warning performance, detecting 480 of 699 events with a detection recall of 0.6867 and 15.714 false positives per 1,000 stable nominal windows. Flow and Feed demonstrated significantly lower event detection recalls of 0.1131 and 0.0711, respectively. Although z-shift achieved an apparent event detection recall of 0.6916, it yielded 712.794 false positives per 1,000 windows; with a false positive rate of approximately 1%, only 1 of 227 events was detected. Therefore, z-offset cannot be considered a reliable prediction target within the current protocol.

Overall, GATE3D should be considered a competitive rather than universally superior prediction system. The strongest and most consistent results were obtained for the hotend, while flow and feed rates were comparable to strong baseline gain metrics, and z-offset remained ineffective. A key limitation is that all experiments used portions of the same Caxton dataset; therefore, the results demonstrate generalization to previously unseen samples, but not yet to other printers, materials, or operating conditions. Future research should focus on independent external validation, target feature enrichment, and adaptive prediction strategies.

5. CONCLUSION

This study developed GATE3D, a hybrid camera-free framework for early anomaly forecasting in FFF/FDM additive manufacturing using only native printer telemetry under a strict leakage-free protocol. On an unseen-print holdout comprising 5,323 windows from 27 previously unseen prints, GATE3D achieved Macro-F1 scores of 0.8866, 0.8830, and 0.9050 for flow rate, feed rate, and hotend, respectively. In contrast, z-offset remained near chance level with a Macro-F1 of 0.5033. Comparative and statistical analysis showed that GATE3D is competitive with strong machine learning and sequence-based models, while a statistically confirmed advantage over HGB was observed only for the hotend (Δ=+0.0035; 95% CI: 0.0006–0.0069). Event-level evaluation further confirmed that the hotend is the most reliable early warning target: 480 out of 699 events were detected, with an event recall rate of 0.6867.

Overall, the results demonstrate that the effectiveness of telemetry-based early anomaly prediction is highly dependent on the target, and that strong window-level performance should be complemented by event-level evaluation. Future work will focus on independent validation across different printers and materials, improved telemetry representation for complex targets such as Z-axis offset, and adaptive prediction strategies for practical FFF/FDM monitoring.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Arailym Kenzhebay	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Ardak Karipzhanova		✓				✓		✓	✓	✓	✓	✓		
Nazira Ospanova	✓		✓	✓			✓			✓	✓		✓	✓
Alibek Barlybayev		✓				✓		✓	✓	✓	✓	✓		
Aigerim Bakiyeva					✓		✓			✓		✓		✓
Anargul Bekenova		✓				✓		✓	✓	✓	✓	✓		
Yerbol Yerbayev					✓		✓			✓		✓		✓
Assemgul Tynykulova					✓		✓			✓		✓		✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY

The data, source code, experimental configurations, reproducible Jupyter notebooks, and supporting results used in this study are publicly available in the GATE3D-FFF-Regime-Forecasting GitHub repository: https://github.com/nazira-o73/GATE3D-FFF-Regime-Forecasting.

REFERENCES

[1] S. C. Daminabo, S. Goel, S. A. Grammatikos, H. Y. Nezhad, and V. K. Thakur, "Fused deposition modeling-based additive manufacturing (3D printing): techniques for polymer material systems," *Materials Today Chemistry*, vol. 16, p. 100248, 2020, doi: 10.1016/j.mtchem.2020.100248.

[2] A. Oleff, B. Küster, M. Stonis, and L. Overmeyer, "Process monitoring for material extrusion additive manufacturing: a state-of-the-art review," *Progress in Additive Manufacturing*, vol. 6, no. 4, pp. 705–730, 2021, doi: 10.1007/s40964-021-00192-4.

[3] G. D. Goh, S. L. Sing, and W. Y. Yeong, "A review on machine learning in 3D printing: applications, potential, and challenges," *Artificial Intelligence Review*, vol. 54, no. 1, pp. 63–94, 2021, doi: 10.1007/s10462-020-09876-9.

[4] Z. Jin, Z. Zhang, and G. X. Gu, "Autonomous in-situ correction of fused deposition modeling printers using computer vision and deep learning," *Manufacturing Letters*, vol. 22, pp. 11–15, 2019, doi: 10.1016/j.mfglet.2019.09.005.

[5] D. A. J. Brion and S. W. Pattinson, "Generalisable 3D printing error detection and correction via multi-head neural networks," *Nature Communications*, vol. 13, no. 1, p. 4654, 2022, doi: 10.1038/s41467-022-31985-y.

[6] M. Islam, M. Bappy, S. Tushar, C. Zamiela, and A. Wasi, "Beyond local receptive fields: vision transformers for real-time surface defect detection in FDM," *The International Journal of Advanced Manufacturing Technology*, vol. 143, no. 9–10, pp. 5399–5420, 2026, doi: 10.1007/s00170-026-17846-8.

[7] U. Delli and S. Chang, "Automated process monitoring in 3D printing using supervised machine learning," *Procedia Manufacturing*, vol. 26, pp. 865–870, 2018, doi: 10.1016/j.promfg.2018.07.111.

[8] J. E. Seppala and K. D. Migler, "Infrared thermography of welding zones produced by polymer extrusion additive manufacturing," *Additive Manufacturing*, vol. 12, pp. 71–76, 2016, doi: 10.1016/j.addma.2016.06.007.

[9] H. Wu, Z. Yu, and Y. Wang, "Real-time FDM machine condition monitoring and diagnosis based on acoustic emission and hidden semi-Markov model," *The International Journal of Advanced Manufacturing Technology*, vol. 90, pp. 2027–2036, 2017, doi: 10.1007/s00170-016-9548-6.

[10] Y. Li, W. Zhao, Q. Li, T. Wang, and G. Wang, "In-situ monitoring and diagnosing for fused filament fabrication process based on vibration sensors," *Sensors*, vol. 19, no. 11, p. 2589, 2019, doi: 10.3390/s19112589.

[11] Z. Li, Z. Zhang, J. Shi, and D. Wu, "Prediction of surface roughness in extrusion-based additive manufacturing with machine learning," *Robotics and Computer-Integrated Manufacturing*, vol. 57, pp. 488–495, 2019, doi: 10.1016/j.rcim.2019.01.004.

[12] Y. Tlegenov, W. F. Lu, and G. S. Hong, "A dynamic model for current-based nozzle condition monitoring in fused deposition modelling," *Progress in Additive Manufacturing*, vol. 4, no. 3, pp. 211–223, 2019, doi: 10.1007/s40964-019-00089-3.

[13] S. Rooney, E. Pitz, and K. Pochiraju, "AutoML-driven diagnostics of the feeder motor in fused filament fabrication machines from direct current signals," *Journal of Intelligent Manufacturing*, vol. 36, no. 3, pp. 1999–2016, 2025, doi: 10.1007/s10845-024-02332-3.

[14] H. Zhou *et al.*, "Informer: beyond efficient transformer for long sequence time-series forecasting," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 12, pp. 11106–11115, 2021, doi: 10.1609/aaai.v35i12.17325.

[15] G. Zerveas, S. Jayaraman, D. Patel, A. Bhamidipaty, and C. Eickhoff, "A transformer-based framework for multivariate time series representation learning," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2021, pp. 2114–2124, doi: 10.1145/3447548.3467401.

[16] S. Tuli, G. Casale, and N. R. Jennings, "TranAD: deep transformer networks for anomaly detection in multivariate time series data," in *Proceedings of the VLDB Endowment*, vol. 15, no. 6, pp. 1201–1214, 2022, doi: 10.14778/3514061.3514067.

[17] S. Kaufman, S. Rosset, C. Perlich, and O. Stitelman, "Leakage in data mining: formulation, detection, and avoidance," *ACM Transactions on Knowledge Discovery from Data*, vol. 6, no. 4, pp. 1–21, 2012, doi: 10.1145/2382577.2382579.

[18] S. Kapoor and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns*, vol. 4, no. 9, p. 100804, 2023, doi: 10.1016/j.patter.2023.100804.

[19] D. R. Roberts *et al.*, "Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure," *Ecography*, vol. 40, no. 8, pp. 913–929, 2017, doi: 10.1111/ecog.02881.

[20] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv Preprint*, 2018, doi: 10.48550/arXiv.1803.01271.

[21] T. Chen and C. Guestrin, "XGBoost: a scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[22] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001, doi: 10.1214/aos/1013203451.

[23] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[24] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, "Measuring and testing dependence by correlation of distances," *The Annals of Statistics*, vol. 35, no. 6, pp. 2769–2794, 2007, doi: 10.1214/009053607000000505.

[25] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. New York, NY, USA: Chapman & Hall/CRC, 1994, doi: 10.1201/9780429246593.

BIOGRAPHIES OF AUTHORS



Arailym Kenzhebay received her bachelor's degree in Information Systems from L.N. Gumilyov Eurasian National University in 2021 and completed her master's degree in Information Systems at the same institution in 2023. Since 2023, she has been pursuing a Ph.D. in Computer Science at Alikhan Bokeikhan University. She is the author of three scientific publications. Her research interests include artificial intelligence, additive manufacturing technologies, and cloud computing. She can be contacted at email: salik2000@icloud.com.



Ardak Karipzhanova holds a Ph.D. in information systems and currently serves as the vice-rector for Information Technology at Alikhan Bokeikhan University. She obtained her bachelor's degree in Mathematics from S.M. Kirov Kazakh State University in 1989, her master's degree in Natural Sciences (Informatics) from the Kazakh Humanitarian-Legal Innovative University in 2015, and her Ph.D. in Information Systems from L.N. Gumilyov Eurasian National University in 2019. She is the author of 21 scientific publications, including articles indexed in Scopus and Web of Science, as well as two scientific monographs and four patents. Her research interests include information security, machine learning, and digital educational technologies. She can be contacted at email: kamilakz2001@mail.ru.



Nazira Ospanova graduated from Shakarim University in 1996. In 2010, she received the academic degree of Candidate of Pedagogical Sciences in Informatics and Education Informatization. She is currently working as a professor in the Department of Information Technologies at Toraighyrov University. She is the author of about 50 publications. Her research interests include issues of informatics, education informatization, and software development in various fields of human activity. She can be contacted at email: ospanova.n@teachers.tou.edu.kz.



Alibek Barlybayev received the B.Eng. degree in information systems from L.N. Gumilyov Eurasian National University, Kazakhstan, in 2009 and the M.S. and Ph.D. degrees in Computer Science from L.N. Gumilyov Eurasian National University, Kazakhstan, in 2011 and 2015, respectively. Currently, he is a director of the Research Institute of Artificial Intelligence, L.N. Gumilyov Eurasian National University. He is also an associate professor of the Department of Artificial Intelligence Technologies, L.N. Gumilyov Eurasian National University, and Higher School of Information Technology and Engineering, Astana International University. His research interests are NLP, the use of neural networks in word processing, smart textbooks, fuzzy logic, stock market price prediction, and information security. He can be contacted at email: frank-ab@mail.ru.



Aigerim Bakiyeva graduated from the Eurasian National University named after L.N. Gumilev with a degree bachelor in Informatica in 2010. In 2019, she defended her dissertation in the specialty 05.13.17 – «Theoretical informatics» and "6D075100 - Informatics, computer engineering and management" and received a candidate of technical sciences and Ph.D. She began her career in 2010 as an assistant teacher at the Department of Social Sciences and Humanities of Kazakh National University of Arts. Currently, she is an assistant professor at the Department of Software Engineering of Astana IT University. She is the author of more than 40 scientific papers, including 2 monographs, and 5 articles in the Scopus database. Scientific interests–information systems, data mining, and natural language processing. She can be contacted at email: aigerimbakiyeva@gmail.com.



Anargul Bekenova master of Technical Sciences, currently working as a lecturer at the West Kazakhstan Agricultural and Technical University named after Zhangir Khan. Her main scientific interests include artificial intelligence, programming, computer mathematics and engineering modeling, adaptive learning in education, management, and information processing. Over the past 5 years, she has published 10 scientific articles, including 8 in journals included in the Scopus database. She is a graduate of the International "Bolashak" program. She can be contacted at email: inabat.77@mail.ru.



Yerbol Yerbayev Ph.D. in the specialty "Electric power Industry", currently, he works at the NJSC «West Kazakhstan Agrarian and Technical University named after Zhangir khan» at the Polytechnic Institute, associate professor. He has more than 20 years of scientific and pedagogical experience and more than 70 scientific papers, including 15 articles in the Scopus database, 10 teaching aids. The Hirsch index is 4. His research interests lie in the fields of electric power engineering, automation, and information technology. He can be contacted at email: erbol.erbaev@mail.ru.



Assemgul Tynykulova has teaching experience at leading universities in Kazakhstan. Her main research interests include optimization methods, decision-making support, and expert systems. Over the past five years, she has published more than ten scientific articles, including four in journals indexed in the Scopus database. She currently holds a Ph.D. degree and works as an acting associate professor at the Higher School of Information Technology and Engineering of Astana International University, Astana. Astana International University, Ph.D., senior lecturer at the Higher School of Information Technology and Engineering, Astana, Kazakhstan. She can be contacted at email: asem_110981@mail.ru.