

Article

Secure Machine Learning Framework for Defect Detection and Quality Enhancement in Injection Molding Processes

Mi Young Kang

Department of Information & Communication Engineering, Honam University, Gwangju 62399, Republic of Korea; kmy2021@honam.ac.kr

Abstract

The Fifth Industrial Revolution (Industry 5.0) requires human-centric mechanisms that preserve the integrity, reproducibility, and interpretability of AI-driven decisions in smart manufacturing. Injection molding generates heterogeneous, imbalanced, and weakly labeled process data, posing reliability and integrity risks to data-driven quality control. This study proposes an integrity-verified and reproducibility-instrumented secure machine learning framework for operating-regime analysis in injection molding that integrates (i) SHA-256-based data-integrity verification at ingestion, (ii) Pearson correlation-based feature selection, and (iii) a Gaussian Mixture Model (GMM) under a passive-adversary threat model with Transport Layer Security (TLS)-secured transmission. Evaluated on real industrial data ($n = 6719$ cycles, seven process variables), correlation-based feature selection retained four non-redundant variables and improved the GMM Silhouette Score from 0.274 ± 0.075 (all features) to 0.323 ± 0.014 (95% CI [0.318, 0.329]), a +18.2% relative improvement (paired $t(29) = 3.39$, $p = 0.002$; Cohen's $d = 0.62$; Wilcoxon $p = 0.022$), while lowering the Davies–Bouldin Index from 1.63 to 1.17. The Silhouette standard deviation of 0.014 over 30 seeds meets the $\sigma \leq 0.02$ reproducibility target. The GMM resolves four interpretable operating regimes—one low-load regime consistent with nominal operation and three elevated-load regimes (left-side, right-side, and bilateral)—with operator-readable per-variable signatures. Relative to hard-partition and projection baselines, the GMM is not Silhouette-optimal but provides an interpretable, generative regime model that meets the $\sigma \leq 0.02$ reproducibility target. The framework operationalizes human-centric manufacturing security as measurable integrity, reproducibility, and interpretability.

Keywords: injection molding; operating-regime analysis; human-centric security; data integrity; Gaussian mixture model; reproducibility; smart manufacturing; Industry 5.0



Academic Editor: Domenico Ursino

Received: 21 May 2026

Revised: 20 June 2026

Accepted: 24 June 2026

Published: 26 June 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The Fifth Industrial Revolution (Industry 5.0) recasts industrial value creation around human–machine collaboration, sustainability, and resilience [1–3]. Within this paradigm, manufacturing security extends beyond perimeter and network defense to the integrity, reproducibility, and interpretability of the AI-driven decision systems on which human operators rely [4,5]. Failures in these properties—silent data corruption, non-reproducible model outputs, or opaque classifications—translate directly into unsafe process adjustments, downstream defects, and erosion of operator trust.

Injection molding, a core process in automotive, electronic, and precision-component manufacturing, illustrates these risks at scale. Modern lines generate large volumes of heterogeneous sensor and PLC data transmitted through industrial APIs to cloud-integrated

Manufacturing Execution Systems (MESs) for quality control [6,7]. Three structural conditions complicate this pipeline: defect samples are rare and class-imbalanced [8,9]; nonlinear process behavior and material variability render fixed-threshold rules brittle [10,11]; and cloud transmission exposes the pipeline to integrity threats such as silent corruption, replay, and man-in-the-middle injection [12,13]. Adversarial machine learning research further shows that small, semantically plausible perturbations can deflect industrial classifiers without operator awareness [14,15].

Recent contributions to the Industry 5.0 security agenda address complementary layers. Nsoh [16] proposed a Human-Centric Zero-Trust Identity Architecture (HC-ZTIA) securing the identity plane. Jho and Youn [17] introduced stateful order-preserving encryption (SOPE) for cloud-resident data confidentiality. Jung and Kim [18] analyzed integrity vulnerabilities in YOLO-based industrial vision pipelines. A complementary gap remains: securing the integrity and reproducibility of unsupervised ML-driven decisions on the shop floor, where labels are scarce and operators act on probabilistic outputs.

1.1. Operational Definition of Human-Centric Manufacturing Security

We operationalize human-centric manufacturing security through three measurable properties of the ML decision pipeline:

- (i) **Data Integrity:** every sensor payload is hashed with SHA-256 in the Device Zone and verified in the Database Zone; the target is a 100% verification pass rate under normal operation, with mismatches triggering quarantine and operator alert.
- (ii) **Decision Reproducibility:** clustering outputs must be statistically stable across random initializations and covariance configurations; the target is a Silhouette Score standard deviation $\sigma \leq 0.02$ over 30 independent runs.
- (iii) **Interpretable Attribution:** defect-related clusters must be associated with human-readable process-variable signatures rather than opaque latent codes.

Throughout this paper, the term “secure” denotes these three measurable properties under the passive-adversary threat model in Section 3.1 rather than resistance to active attacks on the model; the latter is explicitly out of scope (Section 5.3).

1.2. Hypotheses and Contributions

This study tests two hypotheses.

H1. *Integrity-aware preprocessing (hash verification, IQR-based clipping, normalization) and correlation-based feature selection improve clustering reproducibility under industrial data imbalance.*

H2. *Probabilistic soft assignments via GMM provide a generative, interpretable representation of operating regimes—operator-readable per-variable signatures with calibrated per-cycle membership—complementing rather than outperforming hard-assignment baselines on internal cluster-validity indices.*

The main contributions are: (i) a measurable, three-property operationalization of human-centric manufacturing security for unsupervised industrial ML; (ii) an operating-regime-analysis framework integrating SHA-256 integrity verification, TLS-secured transmission, correlation-based feature selection, and GMM clustering, with an explicit passive-adversary threat model; and (iii) empirical evidence on $n = 6719$ real injection molding cycles showing a reproducible Silhouette Score of 0.323 ± 0.014 , a statistically significant +18.2% improvement from feature selection over the all-feature baseline (paired $t(29) = 3.39$, $p = 0.002$), and stable cluster topology across four covariance structures, positioned relative to identity-layer [16], cloud-encryption [17], and pipeline-vulnerability [18] approaches.

2. Related Work

2.1. Human-Centric Security in Industry 5.0

The Industry 5.0 agenda originates from the European Commission's vision of a sustainable, human-centric, resilient European industry [1] and subsequent surveys characterizing the transition from Industry 4.0 [2,3]. Nsoh [16] proposed HC-ZTIA, situating identity as the control plane and extending NIST SP 800-207 [19] zero-trust principles. Jho and Youn [17] proposed SOPE, preserving privacy in cloud databases while supporting range queries. Jung and Kim [18] analyzed integrity vulnerabilities in OpenCV-YOLO-based vision pipelines. Collectively, these address identity, data-at-rest, and pipeline layers; the integrity and reproducibility of unsupervised ML-driven decisions on imbalanced industrial data remain less developed.

2.2. Unsupervised Anomaly Detection for Manufacturing Quality Control

Unsupervised approaches are attractive because defective samples are rare and labels costly [8,20]. GMMs have been widely applied for their probabilistic modeling and interpretable structure [21]. Oluwasegun and Jung [21] applied multivariate GMMs to anomaly detection in nuclear control element drive mechanisms; Geng et al. [22] proposed an unsupervised deep model for online anomaly detection in continuous casting. In injection molding, supervised and ensemble approaches have addressed sink-mark detection [10], shrinkage surrogate modeling [11], and quality prediction [9]. Kim et al. [23] emphasized that preprocessing quality strongly influences anomaly-detection reliability. Most prior works prioritize predictive accuracy and rarely treat reproducibility, integrity, or interpretability as first-class criteria.

2.3. Data Integrity and Security in Industrial Cyber-Physical Systems

Cloud APIs, MESs, and OT networks expose ML pipelines to integrity threats. NIST SP 800-82 Rev. 3 [24] and IEC 62443-3-3 [25] codify integrity, authentication, and access-control requirements. Sebestyen et al. [12] reviewed IoT security and identified data integrity and reliable decision-making as central challenges; Riaz et al. [13] proposed a robust anomaly detector for imbalanced Industrial Internet of Things (IIoT) data; Zhang et al. [26] introduced IDD-Net for data-quality variability. Beyond classical threats, adversarial machine learning shows small perturbations can mislead industrial classifiers [14,15], and federated learning [27] has been proposed as a complementary defense. Recent studies further extend these threats to data poisoning and false-data-injection attacks in IIoT and Industry 5.0 environments [28,29], underscoring the need for integrity verification at ingestion.

2.4. Research Gap and Positioning

Identity-layer (HC-ZTIA [16]) and cryptographic-storage (SOPE [17]) approaches secure who can act and how data are kept at rest but do not constrain the integrity and reproducibility of unsupervised ML-driven decisions on the shop floor; vulnerability analyses [18] reveal attack surfaces but stop short of an affirmative defense. The present study addresses this gap by combining integrity-verified ingestion, TLS-secured transmission, and a reproducibility-evaluated GMM.

Method-selection rationale: The three measurable security properties mentioned in Section 1.1 motivate each methodological choice. Data integrity is enforced by SHA-256 verification and TLS transmission, the minimal mechanisms that detect payload corruption at ingestion under a passive-adversary model. Decision reproducibility and interpretability motivate the choice of a Gaussian Mixture Model over hard-partition or black-box alternatives: its probabilistic soft assignment represents process drift and uncertainty rather than

forcing rigid boundaries, and its per-component means yield operator-readable signatures in physical process variables—both achievable without scarce defect labels. Correlation-based feature selection precedes clustering because the load sensors are strongly collinear ($|\rho| \geq 0.80$), and removing redundancy stabilizes covariance estimation, which is the primary driver of run-to-run reproducibility. These choices are consequences of the security properties the framework is designed to deliver, not arbitrary algorithm selection.

3. Materials and Methods

3.1. Threat Model

We adopt a passive-adversary (honest-but-curious) model augmented with data-integrity threats during transmission. The adversary may observe data in transit between the Device, Database, and Machine Learning Zones and may attempt unauthorized modification (e.g., man-in-the-middle injection, replay).

Active attacks on the ML model itself (adversarial-example crafting, training-data poisoning) and verbatim replay of valid payloads are out of scope for the current evaluation and are identified in Section 5.3 as priority extensions.

3.2. System Architecture

Figure 1 illustrates the architecture, organized as three zones (Device, Database, Machine Learning) connected by a secured Cloud Server. All inter-zone traffic uses TLS 1.3 over HTTPS. In the Device Zone, each acquisition cycle produces a payload to which a SHA-256 hash is appended; in the Database Zone, hash verification gates ingestion; in the Machine Learning Zone, the GMM consumes integrity-verified records and emits cluster outputs with reproducibility metadata (seed, covariance structure, run identifier).

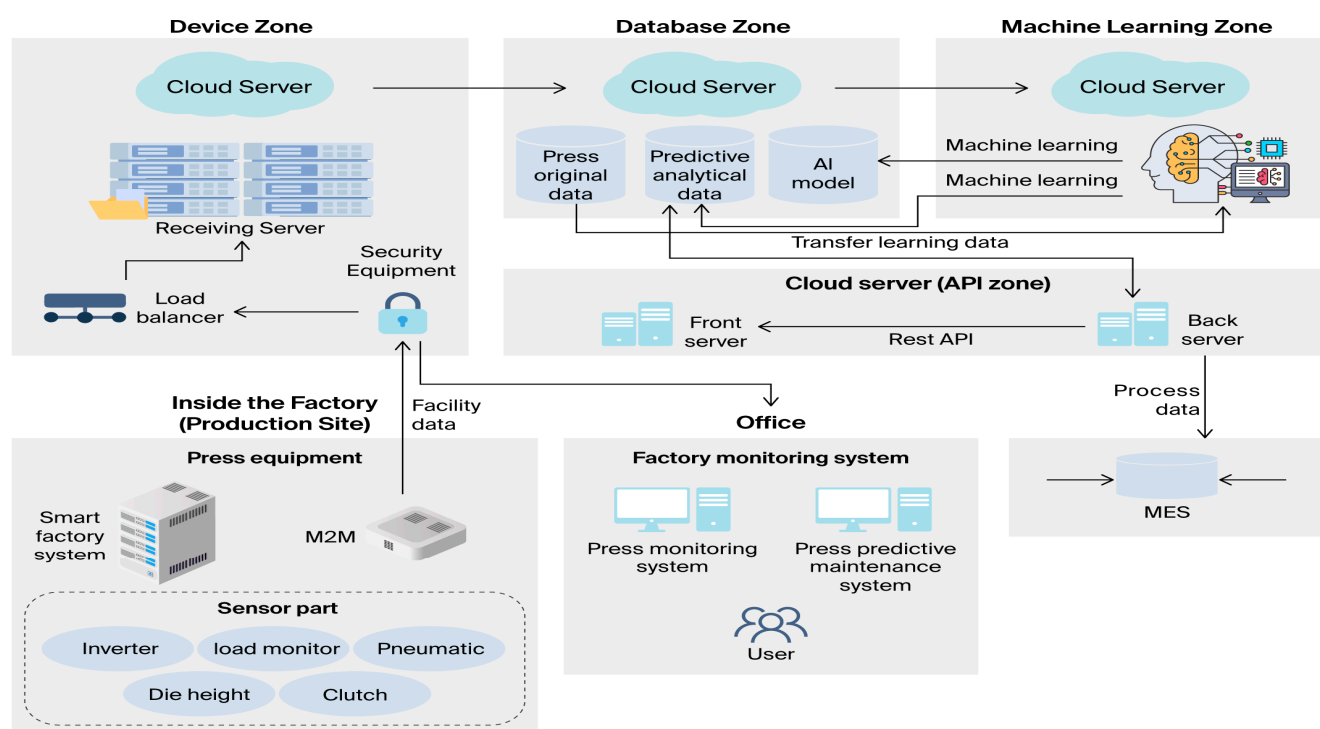


Figure 1. Conceptual architecture of the proposed framework (three zones connected by a secured Cloud Server and integrated with an MES).

3.3. Data Description and Acquisition

Data were collected from an injection molding production line from 25 October 2024 to 26 October 2024, yielding $n = 6719$ cycles. Seven process variables were retained after data-availability filtering: `slide_position`, `load_160_l_rd`, `load_160_l_ton`, `load_160_r_rd`, `load_160_r_ton`, `load_160_lr_rd`, and `load_lr_ton`, acquired from PLCs and the monitoring system. All records were anonymized and consolidated as structured tabular data.

Operator visual-inspection notes were available only at an aggregate level and were not retained as per-cycle labels aligned to the 6719 records; consequently, they are used as qualitative context only and not for any quantitative external validation (Section 5.3). A representative sample of the raw structure is shown in Table 1.

Table 1. Sample of the raw injection molding process data after acquisition from PLCs.

	Working_Time	Slide_Position	Load_160_l_rd	Load_160_l_ton	Load_160_r_rd	Load_160_r_ton	Load_160_lr_rd	Load_lr_ton
0	25 October 2024	476.52	0	0	0	0	0	0
1	25 October 2024	476.52	0	0	0	0	0	0
2	25 October 2024	476.52	0	0	0	0	0	0
3	25 October 2024	476.52	0	0	14	0	14	0
4	25 October 2024	476.52	0	0	14	0	14	0
...	...	...	...	...	...	...	...	...
6714	26 October 2024	477.60	30	0	0	0	30	0
6715	26 October 2024	477.41	30	0	0	0	30	0
6716	26 October 2024	477.42	30	0	0	0	30	0
6717	26 October 2024	477.47	30	0	0	0	30	0
6718	26 October 2024	477.52	30	0	0	0	30	0

3.4. Integrity-Aware Preprocessing

3.4.1. SHA-256 Hash-Based Integrity Verification

For each acquisition cycle c , the Device Zone constructs a payload P_c containing timestamp, machine identifier, and the measured-variable vector. A SHA-256 digest $h_c = \text{SHA-256}(P_c)$ is appended and transmitted under TLS 1.3. In the Database Zone the digest is recomputed and compared bitwise; mismatched records are logged, excluded, and surfaced to operators. In the present dataset, 100% of records passed verification under nominal conditions.

We note that SHA-256 verification detects payload corruption and modification but does not by itself detect a verbatim replay of a valid payload; replay resistance via timestamp-freshness/nonce checking is identified as a delimited extension (Section 5.3). Injected-perturbation evaluation is future work.

3.4.2. Missing-Value Handling, Outlier Treatment, and Normalization

Variables with >10% missing values were excluded; the remaining missing entries were imputed with the per-variable median. Outliers were detected per variable using the IQR criterion (beyond $Q1 - 1.5 \times \text{IQR}$ or $Q3 + 1.5 \times \text{IQR}$) and boundary-clipped rather than removed. All features were then scaled to $[0, 1]$ via min-max normalization.

3.5. Mathematical Formulation of the Gaussian Mixture Model

Notation: the symbols used in Equations (1)–(9) are defined in Table 2.

$$p(x) = \sum_{k=1}^K \pi_k N(x | \mu_k, \Sigma_k) \quad (1)$$

$$N(x | \mu_k, \Sigma_k) = 1 / [(2\pi)^{d/2} |\Sigma_k|^{1/2}] \cdot \exp[-1/2 (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k)] \quad (2)$$

$$\sum_{k=1}^K \pi_k = 1, \pi_k \geq 0 \quad (3)$$

$$r_{ik} = [\pi_k N(x_i | \mu_k, \Sigma_k)] / [\sum_{j=1}^K \pi_j N(x_i | \mu_j, \Sigma_j)] \quad (4)$$

$$N_k = \sum_{i=1}^N r_{ik} \quad (5)$$

$$\pi_k = N_k / N \quad (6)$$

$$\mu_k = (1/N_k) \sum_{i=1}^N r_{ik} \cdot x_i \quad (7)$$

$$\Sigma_k = (1/N_k) \sum_{i=1}^N r_{ik} \cdot (x_i - \mu_k)(x_i - \mu_k)^T \quad (8)$$

$$L(\Theta) = \sum_{i=1}^N \log[\sum_{k=1}^K \pi_k N(x_i | \mu_k, \Sigma_k)] \quad (9)$$

Table 2. Notation used in the GMM/EM formulation.

Symbol	Meaning	Shape
x	process observation vector (one cycle)	R^d
d	number of selected process variables	scalar
K	number of mixture components (=4)	scalar
π_k	mixing weight of component k	scalar, sum = 1
μ_k	mean vector of component k	R^d
Σ_k	covariance matrix of component k	$R^{(d \times d)}$
r_{ik}	responsibility of component k for x_i	scalar in $[0, 1]$
N	number of cycles (=6719)	scalar
N_k	effective count in component k	scalar
Θ	parameter set $\{\pi_k, \mu_k, \Sigma_k\}$	-
ϵ	log-likelihood convergence tolerance (1×10^{-3})	scalar

π_k = mixing weight, μ_k = mean vector, Σ_k = covariance matrix of component k ; r_{ik} = responsibility of component k for observation i ; N_k = effective count in component k ; N = total number of cycles; $\Theta = \{\pi_k, \mu_k, \Sigma_k\}$.

The GMM models the data density as $p(x) = \sum_k \pi_k N(x | \mu_k, \Sigma_k)$ (1), with the multivariate Gaussian (2) subject to $\sum_k \pi_k = 1$, $\pi_k \geq 0$ (3). Parameters Θ are estimated by Expectation Maximization. The E-step computes responsibilities r_{ik} (4); the M-step updates N_k (5), π_k (6), μ_k (7), and Σ_k (8). Iterations proceed until the observed-data log-likelihood $L(\Theta)$ (9) converges within tolerance ϵ .

Equations (1)–(9) are the standard GMM/EM formulation; the contribution is not a modification of them but the integrity-aware pipeline that governs which data enter them and how their outputs are audited. Only SHA-256-verified, preprocessed records (Sections 3.4.1 and 3.4.2) form the input set $\{x_i\}$; the EM updates (Equations (4)–(8)) run under fixed, logged hyperparameters (K , covariance type, tolerance ϵ) over 30 seeds; and the soft assignments r_{ik} (Equation (4)) and log-likelihood $L(\Theta)$ (Equation (9)) are emitted with reproducibility metadata. The configuration choices inside these equations— K , covariance structure, and ϵ —are precisely the knobs whose stability is quantified against the $\sigma \leq 0.02$ target in Section 1.1.

The number of components K was set to $K = 4$ on domain grounds—one low-load regime consistent with nominal operation and three recurring elevated-load regimes (left-side, right-side, and bilateral) identified jointly with line engineers—and verified for stability across seeds and covariance structures (Section 4.6). For transparency, the Bayesian Information Criterion (BIC) for the selected-feature model is reported in Table 3; it decreases monotonically over $K = 2.8$ and does not exhibit an interior minimum or elbow, so BIC was not used as the primary selection criterion. The “full” covariance structure is the primary configuration.

Table 3. BIC over K (full covariance, selected features).

K	2	3	4	5	6	7	8
BIC ($\times 10^3$)	−119.4	−139.1	−154.3	−163.3	−165.3	−175.7	−179.3

3.6. Correlation-Based Feature Selection

Pearson pairwise correlations were computed across the preprocessed variables (Figure 2). Several load-related variables exhibit strong positive correlations, indicating coordinated mechanical behavior, while load_160_r_rd shows weak or negative correlations with the left-side group.

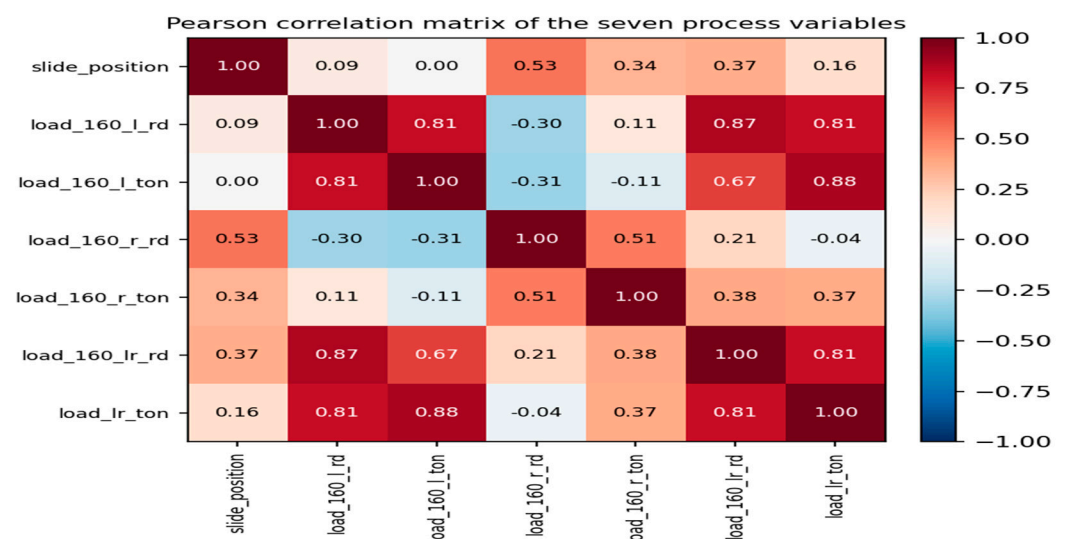


Figure 2. Pearson correlation matrix of the seven process variables.

Applying the $|\rho| \geq 0.80$ criterion, five collinear pairs were identified, all within the left/aggregate load group: load_160_l_rd and load_160_lr_rd ($\rho = 0.87$), load_160_l_ton and load_lr_ton ($\rho = 0.88$), load_160_l_rd and load_160_l_ton ($\rho = 0.81$), load_160_l_rd and load_lr_ton ($\rho = 0.81$), and load_160_lr_rd and load_lr_ton ($\rho = 0.81$). Because min-max normalization equalizes all post-normalization ranges, the redundant variable cannot be chosen by dynamic range; instead, a single representative of the cluster was retained on the basis of partition quality. Each candidate representative (combined with slide_position, load_160_r_rd, and load_160_r_ton) was evaluated over 30 seeds: load_lr_ton, the aggregate left-right tonnage, yielded the most well-separated and reproducible partition (Silhouette 0.323 ± 0.014), compared with 0.292 (load_160_l_ton), 0.214 (load_160_lr_rd), and 0.126 (load_160_l_rd). Accordingly, load_160_l_rd, load_160_l_ton, and load_160_lr_rd were removed. The four retained, non-redundant variables—slide_position, load_160_r_rd, load_160_r_ton, and load_lr_ton—constitute the clustering feature space; all GMM results in Sections 4.2–4.6 are computed on these four variables.

3.7. Implementation, Hyperparameters, Reproducibility, and Complexity

The framework was implemented in Python 3.11 with scikit-learn 1.3.0, NumPy 1.26, and Pandas 2.1. GMM training used $n_components = 4$, $covariance_type = "full"$ (primary), $n_init = 10$, $init_params = "kmeans"$, $max_iter = 200$, $tol = 1 \times 10^{-3}$, $reg_covar = 1 \times 10^{-6}$, with 30 runs at $random_state$ in $\{0, \dots, 29\}$.

Computational complexity: Each EM iteration costs $O(N \times K \times d^2 + K \times d^3)$ (responsibilities and covariance updates plus per-component inversion), giving a total of $O(I \times n_init \times (N \times K \times d^2 + K \times d^3))$ for I iterations and space $O(N \times d + N \times K + K \times d^2)$. With $N = 6719$, $K = 4$, $d = 4$, cost is linear in N ; the measured wall clock was 0.62 s per run ($n_init = 10$) and 0.055 s per single EM run. The GMM scales linearly in N , whereas the One-Class SVM and LOF baselines scale quadratically ($\sim O(N^2)$), a practical deployment advantage. Table 4 summarizes the complexity of all the methods.

Table 4. Computational complexity of the proposed method and baselines.

Method	Time Complexity	Scaling in N
GMM-EM (full covariance)	$O(I \times n_init \times (N \times K \times d^2 + K \times d^3))$	Linear
K-means	$O(I \times n_init \times N \times K \times d)$	Linear
PCA	$O(N \times d^2 + d^3)$	Linear
Isolation Forest	$O(N \times t \times \log \psi)$ scoring	$\sim$ Linear
One-Class SVM (RBF)	$O(N^2) - O(N^3)$	Quadratic+
Local Outlier Factor (LOF)	$O(N^2)$	Quadratic

4. Results

4.1. Process Data Characteristics

Table 5 reports the descriptive statistics of the seven variables ($n = 6719$): `slide_position` has a narrow range (439–491, SD 5.86), while load-related variables span much wider ranges with right-skewed distributions concentrated at zero (Figure 3), indicating that high-load events are rare relative to nominal cycles.

Table 5. Descriptive statistics of the seven process variables ($n = 6719$).

Variable	Mean	Std	Min	25%	50%	75%	Max
<code>slide_position</code>	466.59	5.86	439.03	462.31	463.12	469.09	490.63
<code>load_160_l_rd</code>	12.72	15.10	0.00	0.00	6.00	24.00	67.00
<code>load_160_l_ton</code>	0.43	0.75	0.00	0.00	0.00	1.00	5.00
<code>load_160_r_rd</code>	5.16	7.66	0.00	0.00	0.00	8.00	38.00
<code>load_160_r_ton</code>	0.17	0.38	0.00	0.00	0.00	0.00	2.00
<code>load_160_lr_rd</code>	17.89	14.72	0.00	6.00	15.00	28.00	76.00
<code>load_lr_ton</code>	0.60	0.81	0.00	0.00	0.00	1.00	5.00

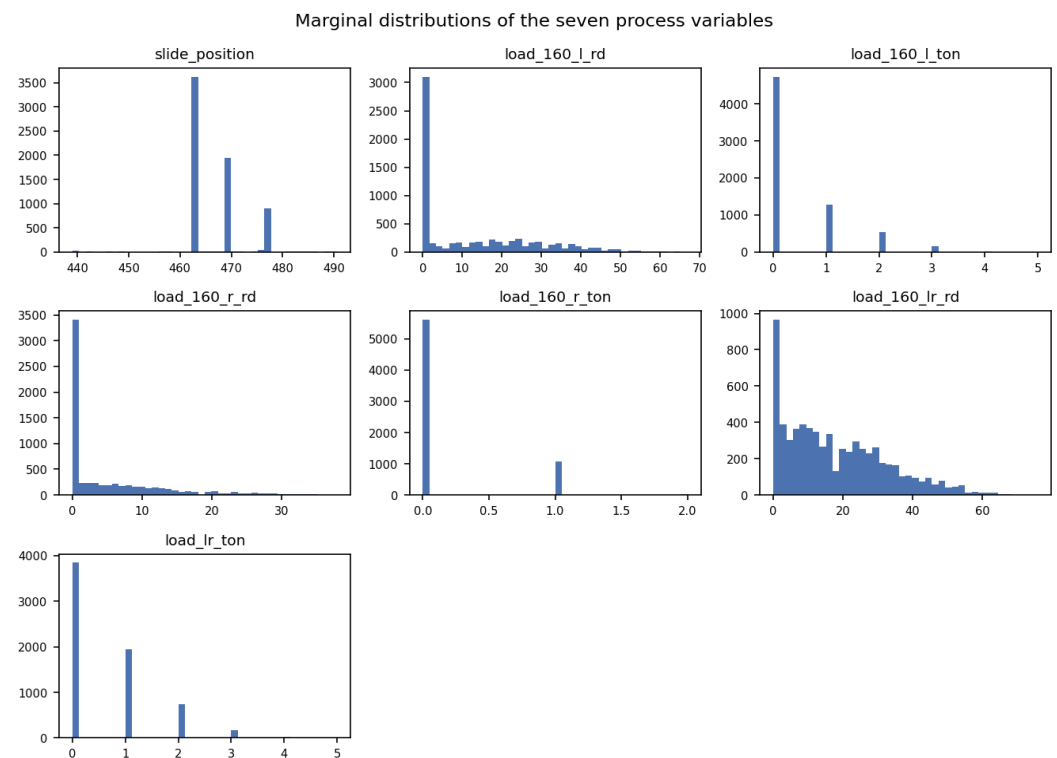


Figure 3. Marginal distributions of the seven process variables (right-skewed, imbalanced).

4.2. GMM-Based Clustering of Operating Regimes

The GMM partitioned the 6719 cycles into four operating regimes of comparable size rather than a single dominant cluster with rare outliers: a low-load regime consistent with nominal operation (N, $n = 1645$), a right-side load regime (R, $n = 2198$), a left-side load regime (L, $n = 1763$), and a bilateral high-load regime (B, $n = 1113$). Because the regimes are comparably represented, they are described as distinct operating regimes rather than a normal/rare-defect dichotomy; the association of the three elevated-load regimes with defect-prone conditions is a domain-based hypothesis not validated against per-cycle operator labels, which were unavailable (Section 3.3). The soft-assignment property represents each cycle by a membership distribution, allowing gradual transitions to be monitored.

4.3. Regime Interpretation

Regime L combines elevated left-side load ($\text{load_160_l_rd} \sim 30.7$, $\text{load_160_lr_rd} \sim 30.7$) with mild left torque and near-zero right-side load, consistent with left-side mechanical loading such as material-flow imbalance or mold misalignment (Figure 4). Regime R shows elevated right-side load ($\text{load_160_r_rd} \sim 8.8$) with negligible left activity. Regime B shows simultaneously elevated left and right load with the highest torque and slide-position offset, consistent with bilateral loading during critical molding stages. Regime N shows uniformly low load and the most stable slide_position . These signatures are operator-readable, satisfying the interpretability property from Section 1.1; they constitute physically motivated hypotheses about defect-prone conditions rather than label-validated defect classes.

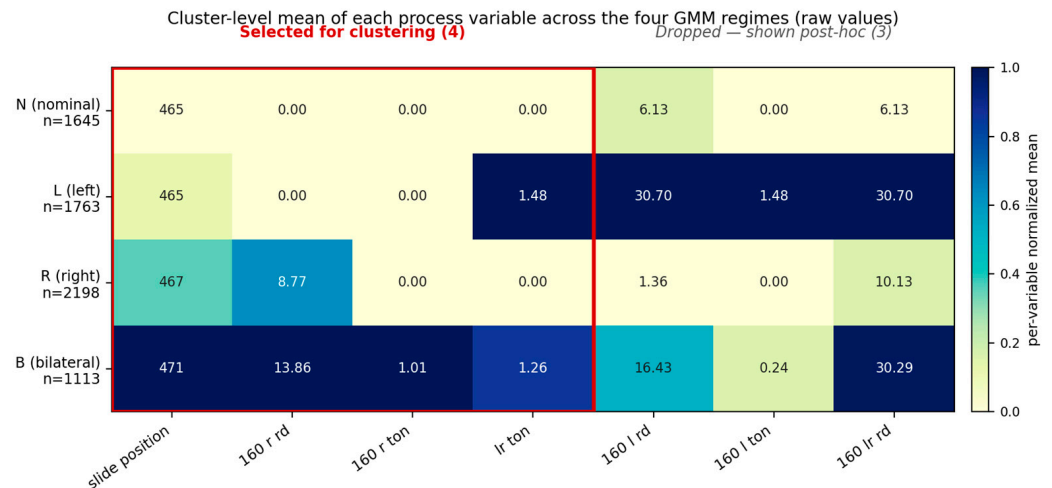


Figure 4. Cluster-level means of each process variable across the four GMM regimes (color-normalized per variable).

Regime occurrence was non-uniform across the two-day window ($n = 930$ on 25 October 2024; $n = 5789$ on 26 October 2024). The bilateral-load regime B was concentrated in the first day's run (586 of 1113 cycles, 53%), whereas the second day was dominated by the low-load and single-side regimes (Figure 5). This non-random temporal concentration supports the interpretation that the regimes reflect genuine process dynamics rather than stochastic noise.

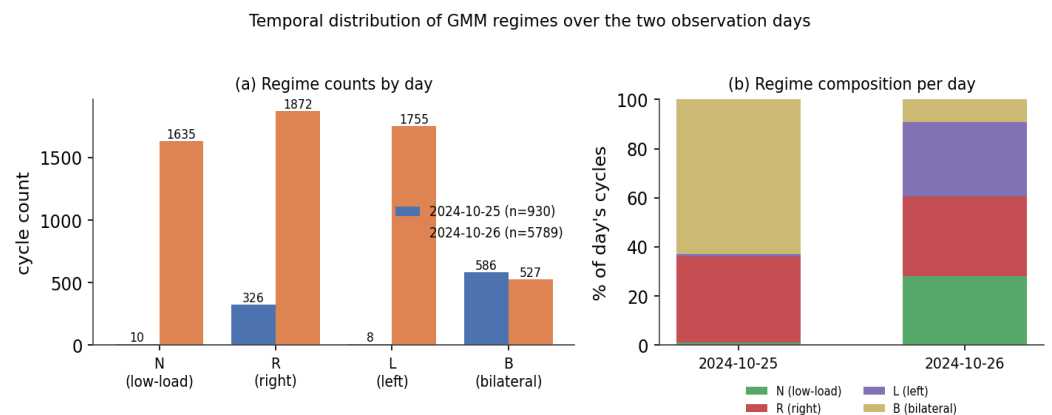


Figure 5. Temporal distribution of the four GMM regimes: (a) counts by day; (b) composition within each day.

4.4. Ablation Study: Contribution of Correlation-Based Feature Selection

Correlation-based feature selection improved the Silhouette Score from 0.274 ± 0.075 (all seven features) to 0.323 ± 0.014 over 30 matched random seeds—a +18.2% relative improvement (paired $t(29) = 3.39$, $p = 0.002$; Cohen's $d = 0.62$; SD of paired differences = 0.080; Wilcoxon $W = 123$, $p = 0.022$). The Davies–Bouldin Index decreased from 1.63 ± 0.24 to 1.17 ± 0.11 , and the Silhouette SD contracted from 0.075 to 0.014, indicating substantially more-reproducible partitions (Table 6).

Table 6. Ablation: correlation-based feature selection.

Metric	Baseline GMM (All 7)	Proposed GMM (4 Selected)	Change
Silhouette (mean $\pm$ SD, $n = 30$)	0.274 ± 0.075	0.323 ± 0.014	+18.2% ($t = 3.39$, $p = 0.002$, $d = 0.62$)
Davies–Bouldin Index (lower better)	1.63 ± 0.24	1.17 ± 0.11	−28%
Silhouette SD (reproducibility)	0.075	0.014	meets $\sigma \leq 0.02$
Wilcoxon signed rank	-	$W = 123$, $p = 0.022$	-

4.5. Comparison with Baselines

The proposed GMM was compared against five unsupervised baselines under identical preprocessing and the same 30-seed schedule: K-means, Principal Component Analysis (PCA) reconstruction-error thresholding, Isolation Forest, One-Class SVM, and LOF (Table 7). On the Silhouette criterion, the hard-partition baseline (K-means) and reconstruction baselines score higher than the proposed GMM. The Silhouette rewards compact, well-separated hard partitions and does not measure the properties this framework targets. The GMM is therefore positioned not as the Silhouette-optimal or the most seed-stable method—K-means (0.476 ± 0.001) and Isolation Forest (0.352 ± 0.011) are both more reproducible on this dataset—but as a generative model that yields operator-readable per-variable regime signatures while still meeting the absolute reproducibility target ($\sigma = 0.014 \leq 0.02$). On this near-deterministic data, the soft assignments are sharp (mean maximum posterior responsibility 0.999; only 2 of 6719 cycles below 0.90), so their practical value lies in monitoring drift toward regime boundaries rather than in resolving ambiguous memberships. The full comparison is reported transparently.

Table 7. Comparison against unsupervised baselines (30-seed schedule).

Method	Silhouette (Mean $\pm$ SD)	DBI
K-means ($k = 4$)	0.476 ± 0.001	0.80
PCA (95th pct)	0.398 ± 0.000	0.76
Isolation Forest	0.352 ± 0.011	1.55
One-Class SVM	0.323 ± 0.000	3.17
LOF	0.051 ± 0.000	4.67
GMM (proposed)	0.323 ± 0.014	1.17

4.6. Reproducibility and Stability

Over 30 random seeds, the proposed (full-covariance, four-feature) model attained a Silhouette standard deviation of 0.014, satisfying the $\sigma \leq 0.02$ target. Across covariance structures, the four-regime topology was preserved (Figure 6), with 30-seed mean Silhouette Scores of 0.352 ± 0.016 (tied), 0.326 ± 0.021 (diagonal), 0.414 ± 0.049 (spherical), and 0.323 ± 0.014 (full). The primary full and tied configurations met the $\sigma \leq 0.02$ target; the spherical structure showed higher run-to-run variability (SD 0.049), reported for completeness.

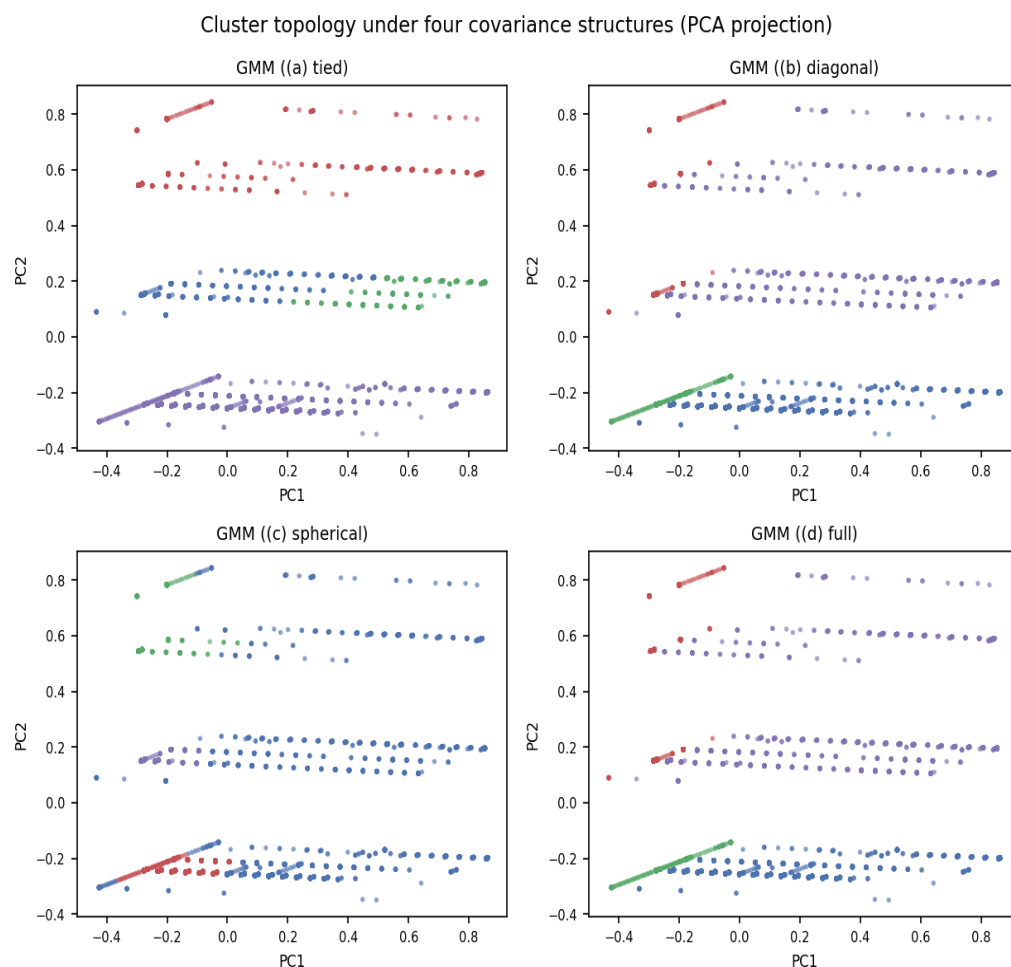


Figure 6. Cluster topologies under four covariance structures (PCA projection).

To assess the choice of K , Silhouette, Davies–Bouldin, and BIC were computed across $K = 2$ – 8 on the selected features (Table 8). No internal index selects $K = 4$: Silhouette is maximized and Davies–Bouldin minimized at $K = 3$, while BIC decreases monotonically; $K = 4$ is therefore a domain-informed choice (Section 5.3).

Table 8. Silhouette, Davies–Bouldin Index, and BIC across the number of components K (selected four features, 30 seeds).

K	Silhouette	Davies–Bouldin	BIC ($\times 10^3$)
2	0.367 ± 0.033	1.354 ± 0.037	−119.4
3	0.392 ± 0.008	1.141 ± 0.143	−139.1
4	0.323 ± 0.014	1.174 ± 0.110	−154.3
5	0.293 ± 0.018	1.307 ± 0.069	−163.3
6	0.319 ± 0.023	1.239 ± 0.154	−165.3
7	0.370 ± 0.040	1.236 ± 0.115	−175.7
8	0.387 ± 0.030	1.444 ± 0.831	−179.3

5. Discussion

The results support a revised reading of the two hypotheses. Integrity-aware preprocessing with correlation-based feature selection (H1) produced a statistically significant, moderate improvement in clustering quality (Silhouette +18.2%, $p = 0.002$, Cohen’s $d = 0.62$)

and a marked reduction in run-to-run variance (SD 0.075 $\rightarrow$ 0.014), supporting the reproducibility property. The probabilistic soft-assignment property of the GMM (H2) yielded interpretable, operator-readable regime signatures. We note explicitly that the GMM does not achieve the highest Silhouette among the evaluated methods; its contribution lies in a generative, interpretable representation of operating regimes—operator-readable per-variable signatures with calibrated membership—while meeting the absolute $\sigma \leq 0.02$ reproducibility target, rather than in maximizing a single internal index or in being the most seed-stable method.

5.1. Positioning Within the Industry 5.0 Security Stack

The framework complements rather than replaces recent work: HC-ZTIA [16] secures identity, SOPE [17] secures data-at-rest confidentiality, and the OpenCV-YOLO analysis [18] surfaces pipeline vulnerabilities. The present work addresses the integrity and reproducibility of unsupervised ML-driven decisions during process execution, pointing toward a defense-in-depth view of human-centric security in Industry 5.0.

5.2. Practical Implications for Operators

The soft-assignment property lets operators monitor the confidence margin between competing regimes, enabling earlier attention to cycles drifting from the low-load state. Regime signatures are expressed in physical process variables, so corrective actions can be reasoned about without algorithmic expertise.

5.3. Limitations

Several limitations should be acknowledged. First, $K = 4$ was set on domain grounds (one nominal and three elevated-load regimes identified with line engineers) and reflects the present line and product mix; internal indices over $K = 2\text{--}8$ on the selected features do not exhibit an interior optimum at $K = 4$ (Silhouette and Davies–Bouldin favor $K = 3$, and BIC decreases monotonically), so $K = 4$ is a domain-informed modeling decision rather than one selected by an internal criterion. Second, per-cycle operator labels were unavailable, so the regimes are not validated against named defect categories; a label-based external evaluation is future work. Third, the evaluation uses a single injection molding line over a two-day window with uneven daily volume ($n = 930$ on 25 October 2024; $n = 5789$ on 26 October 2024); the results should be read as a single-line feasibility demonstration of the integrity-, reproducibility-, and interpretability-oriented pipeline rather than as evidence of cross-line or cross-material generalization. Fourth, the threat model addresses passive adversaries and transmission-integrity threats; active attacks on the model (poisoning, evasion) and verbatim replay are out of scope. Fifth, privacy-preserving techniques such as differential privacy and federated learning [27] are not yet integrated.

5.4. Future Work

Future work will (i) conduct multi-line, multi-shift, and multi-material validation as the primary objective; (ii) add per-cycle labeled external validation (precision/recall/F1, confusion matrix, Adjusted Rand Index); (iii) extend the threat model and evaluation to adversarial robustness, data-poisoning resistance, and replay resistance [28,29]; (iv) incorporate online concept-drift detection so that K and the component parameters adapt as the product mix evolves [30–32]; and (v) integrate federated, privacy-preserving learning for multi-factory deployment.

6. Conclusions

This study proposed a secure machine learning framework—operationalizing human-centric manufacturing security as measurable integrity, reproducibility, and interpretability—for operating-regime analysis in injection molding under an explicit passive-adversary threat model. Integrating SHA-256 ingestion verification, TLS-secured transmission, correlation-based feature selection, and GMM clustering, the framework attained a Silhouette Score of 0.323 ± 0.014 over 30 seeds (95% CI [0.318, 0.329]), a statistically significant +18.2% improvement from feature selection over the all-feature baseline ($p = 0.002$), with the Davies–Bouldin Index reduced from 1.63 to 1.17 and decision reproducibility meeting the $\sigma \leq 0.02$ target. Rather than outperforming hard-partition baselines on internal indices, the framework contributes integrity-verified ingestion, reproducible probabilistic clustering, and interpretable attribution to the emerging Industry 5.0 security stack. Future work will extend validation to multiple lines and materials and to active adversarial threats.

Funding: This study was supported by research funds from Honam University, 2023.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to industrial confidentiality agreements.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. European Commission. *Industry 5.0: Towards a Sustainable, Human-Centric and Resilient European Industry*; Publications Office of the EU: Luxembourg, 2021.
2. Xu, X.; Lu, Y.; Vogel-Heuser, B.; Wang, L. Industry 4.0 and Industry 5.0-Inception, conception and perception. *J. Manuf. Syst.* **2021**, *61*, 530–535. [\[CrossRef\]](#)
3. Leng, J.; Sha, W.; Wang, B.; Zheng, P.; Zhuang, C.; Liu, Q.; Wuest, T.; Mourtzis, D.; Wang, L. Industry 5.0: Prospect and retrospect. *J. Manuf. Syst.* **2022**, *65*, 279–295. [\[CrossRef\]](#)
4. Maddikunta, P.K.R.; Pham, Q.-V.; Prabadevi, B.; Deepa, N.; Dev, K.; Gadekallu, T.R.; Ruby, R.; Liyanage, M. Industry 5.0: A survey on enabling technologies and potential applications. *J. Ind. Inf. Integr.* **2022**, *26*, 100257. [\[CrossRef\]](#)
5. Folgado, F.J.; Calderon, D.; Gonzalez, I.; Calderon, A.J. Review of Industry 4.0 from the Perspective of Automation and Supervision Systems. *Electronics* **2024**, *13*, 782. [\[CrossRef\]](#)
6. Tao, F.; Qi, Q.; Liu, A.; Kusiak, A. Data-Driven Smart Manufacturing. *J. Manuf. Syst.* **2018**, *48*, 157–169. [\[CrossRef\]](#)
7. Soori, M.; Arezoo, B.; Dastres, R. Internet of Things for Smart Factories in Industry 4.0, a Review. *Internet Things Cyber-Phys. Syst.* **2023**, *3*, 192–204. [\[CrossRef\]](#)
8. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of Machine Learning to Machine Fault Diagnosis: A Review and Roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587. [\[CrossRef\]](#)
9. Jung, H.; Jeon, J.; Choi, D.; Park, J.Y. Application of Machine Learning Techniques in Injection Molding Quality Prediction: Implications on Sustainable Manufacturing Industry. *Sustainability* **2021**, *13*, 4120. [\[CrossRef\]](#)
10. Obregon, J.; Hong, J.; Jung, J.-Y. Rule-Based Explanations Based on Ensemble Machine Learning for Detecting Sink Mark Defects in the Injection Moulding Process. *J. Manuf. Syst.* **2021**, *60*, 392–405. [\[CrossRef\]](#)
11. Wenzel, M.; Raisch, S.R.; Schmitz, M.; Hopmann, C. Comparison of Hybrid Machine Learning Approaches for Surrogate Modeling Part Shrinkage in Injection Molding. *Polymers* **2024**, *16*, 2465. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Sebestyen, H.; Popescu, D.E.; Zmaranda, R.D. A Literature Review on Security in the Internet of Things. *Computers* **2025**, *14*, 61. [\[CrossRef\]](#)
13. Riaz, R.; Han, G.; Shaukat, K.; Ullah Khan, N.; Zhu, H. A Robust Anomaly Detector for Imbalanced Industrial Internet of Things Data. *J. Comput. Des. Eng.* **2025**, *12*, 46–60. [\[CrossRef\]](#)
14. Goodfellow, I.J.; Shlens, J.; Szegedy, C. Explaining and Harnessing Adversarial Examples. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
15. Madry, A.; Makelov, A.; Schmidt, L.; Tsipras, D.; Vladu, A. Towards Deep Learning Models Resistant to Adversarial Attacks. In Proceedings of the International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 30 April–3 May 2018.
16. Nsoh, J.T. Human-Centric Zero Trust Identity Architecture for the Fifth Industrial Revolution: A JEPA-Driven Approach to Adaptive Identity Governance. *Electronics* **2026**, *15*, 1878. [\[CrossRef\]](#)
17. Jho, N.-S.; Youn, T.-Y. Stateful Order-Preserving Encryption for Secure Cloud Databases. *Electronics* **2026**, *15*, 1412. [\[CrossRef\]](#)

18. Jung, D.-Y.; Kim, N.-H. Analysis of OpenCV Security Vulnerabilities in YOLO v10-Based IP Camera Image Processing Systems for Disaster Safety Management. *Electronics* **2025**, *14*, 3216. [\[CrossRef\]](#)
19. NIST SP 800-207; Zero Trust Architecture. NIST: Gaithersburg, MD, USA, 2020. [\[CrossRef\]](#)
20. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly Detection: A Survey. *ACM Comput. Surv.* **2009**, *41*, 1–58. [\[CrossRef\]](#)
21. Oluwasegun, A.; Jung, J.-C. A Multivariate Gaussian Mixture Model for Anomaly Detection in Transient Current Signature of Control Element Drive Mechanism. *Nucl. Eng. Des.* **2023**, *402*, 112098. [\[CrossRef\]](#)
22. Geng, M.-Y.; Li, Z.-Y.; Xu, Y.-H.; Liu, S.-L.; Ai, Y.-B.; Zhang, W.-D. An Unsupervised Deep Learning-Based Online Anomaly Detection Model for Mold Level in Continuous Casting Process. *J. Iron Steel Res. Int.* **2026**, *33*, 59. [\[CrossRef\]](#)
23. Kim, S.; Seo, H.; Lee, E.C. Advanced Anomaly Detection in Manufacturing Processes: Leveraging Feature Value Analysis for Normalizing Anomalous Data. *Electronics* **2024**, *13*, 1384. [\[CrossRef\]](#)
24. NIST SP 800-82 Rev. 3; Guide to Operational Technology (OT) Security. NIST: Gaithersburg, MD, USA, 2023. [\[CrossRef\]](#)
25. IEC 62443-3-3:2013; Industrial Communication Networks-Network and System Security-Part 3-3. IEC: Geneva, Switzerland, 2013.
26. Zhang, Z.; Zhou, M.; Wan, H.; Li, M.; Li, G.; Han, D. IDD-Net: Industrial Defect Detection Method Based on Deep-Learning. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106390. [\[CrossRef\]](#)
27. McMahan, H.B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A. Communication-Efficient Learning of Deep Networks from Decentralized Data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017, Fort Lauderdale, FL, USA, 20–22 April 2017; pp. 1273–1282.
28. Firouzi, A.; Dadkhah, S.; Maret, S.A.; Ghorbani, A.A. DataSense: A Real-Time Sensor-Based Benchmark Dataset for Attack Analysis in IIoT with Multi-Objective Feature Selection. *Electronics* **2025**, *14*, 4095. [\[CrossRef\]](#)
29. Habib, A.A.; Hasan, M.K.; Hassan, R.; Islam, S.; Abbas, H.S. False Data Injection Attack Dataset for Classification, Identification, and Detection for IIoT in Industry 5.0. *Data Brief* **2025**, *61*, 111692. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Chapelin, J.; Voisin, A.; Rose, B.; Iung, B.; Steck, L.; Chaves, L.; Lauer, M.; Jotz, O. Data-Driven Drift Detection and Diagnosis Framework for Predictive Maintenance of Heterogeneous Production Processes: Application to a Multiple Tapping Process. *Eng. Appl. Artif. Intell.* **2025**, *139*, 109552. [\[CrossRef\]](#)
31. Hovakimyan, G.; Bravo, J.M. Evolving Strategies in Machine Learning: A Systematic Review of Concept Drift Detection. *Information* **2024**, *15*, 786. [\[CrossRef\]](#)
32. Hinder, F.; Vaquet, V.; Hammer, B. One or Two Things We Know about Concept Drift—A Survey on Monitoring in Evolving Environments. Part A: Detecting Concept Drift. *Front. Artif. Intell.* **2024**, *7*, 1330257. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.