Article

Not peer-reviewed version

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

Mi-Young Kang *

Posted Date: 29 December 2025

doi: 10.20944/preprints202512.2583.v1

Keywords: injection molding; defect detection; machine learning; manufacturing security; data integrity; Gaussian Mixture Model; quality enhancement; smart manufacturing doi:10.20944/preprints202512.2583.v1

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

Injection molding processes generate large volumes of heterogeneous process data that reflect complex and dynamic manufacturing conditions, directly influencing product quality. Conventional quality control approaches based on fixed statistical thresholds and operator experience often fail to ensure stable and secure operation under such variability, leading to increased defect rates and reduced process reliability. Moreover, insufficient consideration of data integrity and process stability can compromise the robustness of data-driven manufacturing systems. This study proposes a **secure machine learning framework** for defect detection and quality enhancement in injection molding processes. The framework integrates systematic data preprocessing, correlation analysis, and an unsupervised learning approach based on a **Gaussian Mixture Model (GMM)** to achieve reliable defect classification. Key process parameters contributing to quality deviations are identified through statistical correlation analysis, and process data are clustered into one normal group and three abnormal groups corresponding to different defect types without requiring labeled datasets. By emphasizing data integrity, stable clustering behavior, and resilience to process variability, the proposed framework enhances manufacturing security and reduces the risk of misclassification. Experimental results using real injection molding process data demonstrate effective defect reduction, improved process parameter optimization, and enhanced operational efficiency. By enhancing defect detectability and reducing misclassification risks, the proposed framework supports safer and more human-centric manufacturing environments by improving operator trust, decision confidence, and preventive intervention capability.

Keywords: injection molding; defect detection; machine learning; manufacturing security; data integrity; Gaussian Mixture Model; quality enhancement; smart manufacturing

1. Introduction

Injection molding is a core manufacturing process widely used in industries such as automotive, electronics, and precision components, where product quality, process stability, and cost efficiency are critical to operational competitiveness. With the rapid adoption of smart manufacturing and Industry 4.0, modern injection molding systems generate large volumes of heterogeneous data from sensors, controllers, and monitoring equipment, enabling data-driven quality control and defect prevention [1–3]. However, the intrinsic variability and complexity of molding conditions—associated with nonlinear process behavior, fluctuating material states, and dynamic machine responses—often challenge conventional quality management practices that rely on fixed statistical thresholds and operator experience [4–6]. These limitations can degrade defect detection accuracy and hinder consistent process control, ultimately undermining reliable decision-making in highly automated production environments.

Recent studies have increasingly explored machine learning-based approaches for defect detection and quality control in manufacturing processes, including injection molding [7–9].

Although supervised learning models have demonstrated promising predictive performance, they typically require large volumes of labeled defect data, which are costly to obtain and often impractical in industrial settings where defective samples are scarce and class distributions are highly imbalanced [10,11]. Moreover, injection molding process data commonly exhibit non-normal distributions, nonlinear relationships among parameters, and sensitivity to internal and external disturbances, which can reduce model generalizability and lead to unstable decisions when deployment conditions deviate from training environments [12,13]. These challenges highlight the need for robust data-driven approaches that can operate effectively under limited labeling conditions while maintaining stable, reliable, and integrity-preserving decision-making in real-world production systems.

To address these challenges, this study proposes a **secure machine learning framework** for defect detection and quality enhancement in injection molding processes. The proposed framework integrates systematic data preprocessing and correlation analysis to identify critical process parameters contributing to quality deviations and employs an unsupervised **Gaussian Mixture Model (GMM)** to cluster process data into one normal group and multiple abnormal groups, enabling defect classification without requiring labeled datasets [14,15]. The main contributions of this study are as follows: (i) development of a data-driven, security-aware defect detection framework that strengthens process reliability, data integrity, and stable decision-making in injection molding systems [16–18]; (ii) demonstration of the effectiveness of GMM-based clustering in distinguishing normal conditions and multiple defect patterns under nonlinear and non-normal manufacturing data distributions; and (iii) experimental validation using real injection molding process data, confirming that the proposed approach supports defect reduction, process parameter optimization, and improved operational efficiency. In addition, by emphasizing stable decision-making and interpretable defect patterns, the proposed framework contributes to human-centric industrial security, enabling operators to better understand process anomalies and respond proactively to potential quality risks. By combining machine learning-based defect detection with manufacturing security considerations, this work contributes to the advancement of reliable and secure smart manufacturing systems aligned with Industry 4.0.

2. Materials and Methods

This section describes the materials, data sources, and methodological procedures employed to develop and validate the proposed secure machine learning framework for defect detection in injection molding processes. The overall methodology follows a data-driven workflow that integrates process data acquisition, data preprocessing, statistical correlation analysis, and unsupervised machine learning–based defect classification. Real injection molding process data collected from industrial equipment were used to ensure practical relevance and reproducibility. To enhance the reliability and stability of defect detection, particular attention was given to data quality management, including handling missing values, outliers, and data imbalance, as well as ensuring consistent parameter scaling. Based on the processed dataset, a Gaussian Mixture Model (GMM) was constructed to cluster process data into normal and abnormal patterns, forming the basis for secure and robust defect identification under realistic manufacturing conditions. Figure 1 illustrates the overall workflow of the proposed secure machine learning framework, including process data acquisition, data preprocessing, correlation analysis, and GMM-based defect detection.

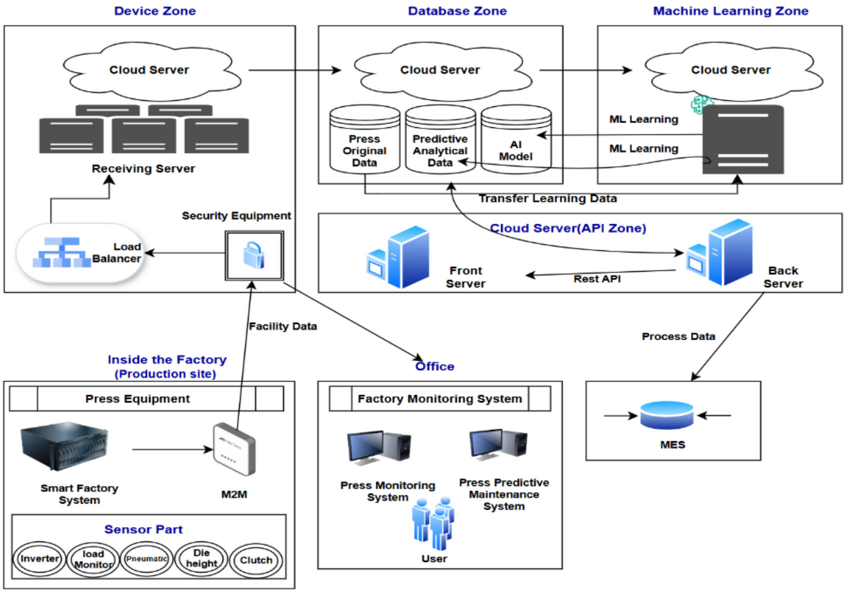


Figure 1. Overall workflow of the proposed secure machine learning framework for defect detection and quality enhancement in injection molding processes.

2.1. Data Description and Acquisition

The dataset used in this study was collected from an industrial injection molding process operated under real production conditions. Process data were acquired directly from programmable logic controllers (PLCs) and associated monitoring systems installed on injection molding equipment. These data represent key physical and operational states of the molding process, including machine motion, load conditions, and process response variables recorded during each production cycle. The collected dataset reflects both normal operating conditions and defective production states, providing a realistic basis for data-driven defect analysis.

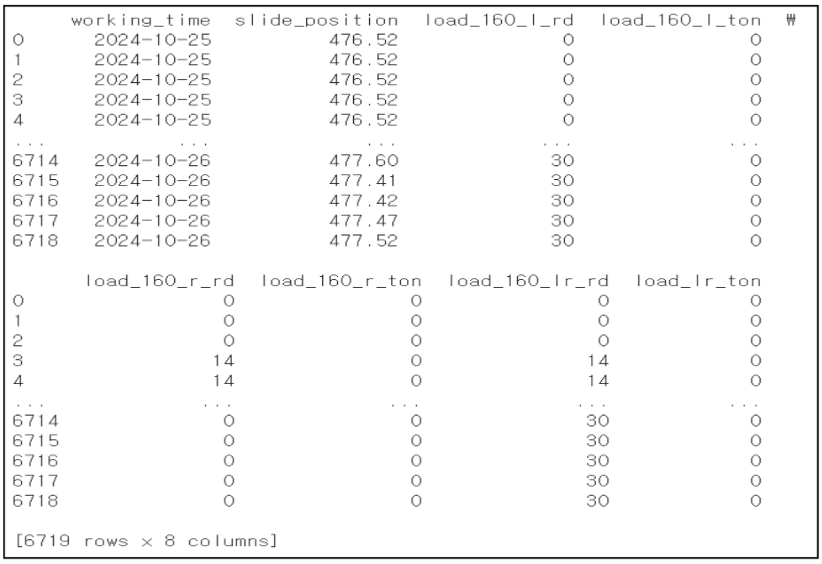
To ensure data reliability and consistency, process variables were sampled at fixed intervals synchronized with the molding cycle. The dataset includes multiple continuous variables related to machine movement and load characteristics, which are known to have a direct influence on product quality in injection molding processes. Visual inspection results performed by experienced operators were used as reference information for identifying defective production cycles, although these labels were not directly employed in the unsupervised learning stage. This approach allows the proposed framework to remain applicable in environments where labeled defect data are limited or partially unavailable.

All collected data were anonymized and organized into structured tabular form prior to analysis. The resulting dataset captures the inherent variability, nonlinearity, and imbalance commonly observed in industrial manufacturing data, thereby providing a suitable foundation for evaluating the robustness and practical applicability of the proposed secure machine learning framework.

2.2. Data Description and Acquisition

Data preprocessing is a critical step in ensuring the reliability and robustness of machine learning–based defect detection in injection molding processes. Industrial manufacturing data often contain missing values, outliers, noise, and inconsistencies caused by sensor faults, communication delays, or unstable operating conditions. If not properly handled, these issues can significantly degrade model performance and lead to unstable or misleading defect classification results. Therefore, systematic data preprocessing and quality management were applied prior to model construction.

First, missing values were identified and treated based on their occurrence rate and relevance to key process parameters. Variables with excessive missing values were excluded from further analysis, while minor missing entries were handled using appropriate imputation strategies to preserve data continuity. Second, outliers were detected using statistical criteria based on the interquartile range (IQR) to identify abnormal observations that deviated significantly from normal operating ranges. Depending on their impact, outliers were either removed or adjusted to boundary values to reduce distortion while maintaining the underlying data distribution. The effects of missing value treatment, outlier handling, and normalization are summarized in Figure 2.



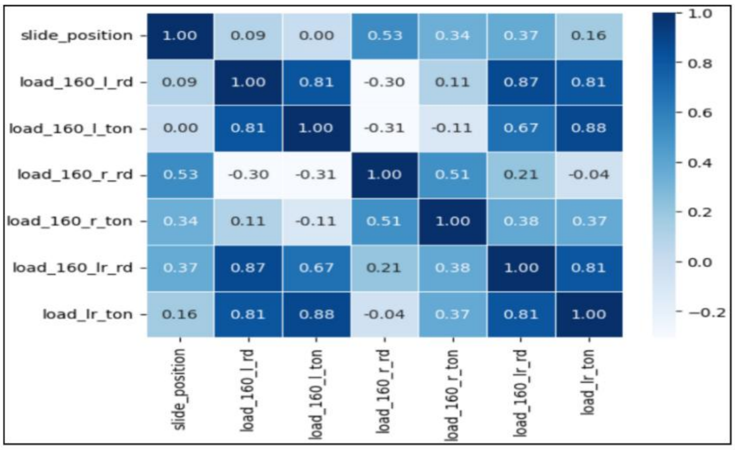
```
     working_time  slide_position  load_160_l_rd  load_160_l_ton  ₩
0      2024-10-25          476.52              0               0
1      2024-10-25          476.52              0               0
2      2024-10-25          476.52              0               0
3      2024-10-25          476.52              0               0
4      2024-10-25          476.52              0               0
...           ...             ...            ...             ...
6714   2024-10-26          477.60             30               0
6715   2024-10-26          477.41             30               0
6716   2024-10-26          477.42             30               0
6717   2024-10-26          477.47             30               0
6718   2024-10-26          477.52             30               0

      load_160_r_rd  load_160_r_ton  load_160_lr_rd  load_lr_ton
0                 0               0               0            0
1                 0               0               0            0
2                 0               0               0            0
3                14               0              14            0
4                14               0              14            0
...             ...             ...             ...          ...
6714              0               0              30            0
6715              0               0              30            0
6716              0               0              30            0
6717              0               0              30            0
6718              0               0              30            0

[6719 rows x 8 columns]
```

Figure 2. Results of data preprocessing, including missing value handling, outlier treatment, and normalization to ensure data integrity and reliability.

To ensure consistent scaling across heterogeneous process variables, normalization was applied to all continuous features. This step prevents dominance of high-magnitude variables during distance-based clustering and contributes to stable model convergence. In addition, basic data quality indicators—including completeness, consistency, and validity—were evaluated to confirm that the processed dataset accurately represents real production conditions. Through these preprocessing procedures, the dataset was transformed into a reliable and integrity-preserving input suitable for secure and stable defect detection in subsequent analysis stages.

2.3. Correlation Analysis

Correlation analysis was conducted to examine the relationships between process variables and to identify key parameters that contribute to quality deviations in injection molding processes. Understanding inter-variable dependencies is essential for reducing model redundancy, improving interpretability, and establishing a reliable foundation for data-driven defect detection. Based on the preprocessed dataset, pairwise correlation coefficients were calculated for all continuous process variables to evaluate the strength and direction of linear relationships. The correlation relationships among process variables are visualized in Figure 3.



	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
slide_position	1.00	0.09	0.00	0.53	0.34	0.37	0.16
load_160_l_rd	0.09	1.00	0.81	-0.30	0.11	0.87	0.81
load_160_l_ton	0.00	0.81	1.00	-0.31	-0.11	0.67	0.88
load_160_r_rd	0.53	-0.30	-0.31	1.00	0.51	0.21	-0.04
load_160_r_ton	0.34	0.11	-0.11	0.51	1.00	0.38	0.37
load_160_lr_rd	0.37	0.87	0.67	0.21	0.38	1.00	0.81
load_lr_ton	0.16	0.81	0.88	-0.04	0.37	0.81	1.00

Figure 3. Correlation matrix of selected process variables, highlighting inter-variable dependencies relevant to defect formation.

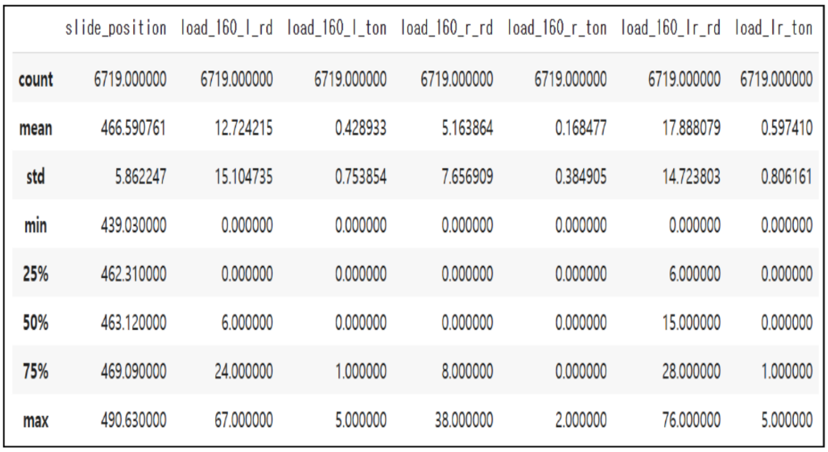
The analysis revealed that several load-related variables exhibited strong positive correlations, indicating coordinated behavior under normal operating conditions. Such variables were considered representative indicators of process stability and mechanical response. In contrast, weak or negative correlations were observed between certain motion-related and load-related parameters, suggesting sensitivity to abnormal process states and potential defect formation. These correlation patterns provide valuable insights into how process dynamics evolve under different operating conditions.

Variables with high correlation coefficients were carefully assessed to avoid multicollinearity effects in subsequent analysis, while parameters demonstrating meaningful associations with quality deviations were retained for further modeling. The results of the correlation analysis were therefore used not only to reduce dimensionality but also to guide the selection of informative features for the unsupervised learning stage. By integrating correlation-based feature assessment, the proposed framework enhances model robustness and supports stable and reliable defect clustering in the presence of complex and nonlinear manufacturing data.

3. Results

3.1. Clustering Results and Interpretation

The clustering results obtained using the Gaussian Mixture Model (GMM) provide clear insights into the underlying structure of the injection molding process data. Based on the predefined number of components, the GMM partitioned the dataset into four distinct clusters, consisting of one dominant cluster representing normal operating conditions and three smaller clusters corresponding to different abnormal or defect-related states. The normal cluster accounted for the majority of samples, while the three defect-related clusters collectively represented a relatively small portion of the dataset, reflecting typical industrial data imbalance. This distribution aligns with the inherent imbalance commonly observed in industrial manufacturing data, where normal production cycles significantly outnumber defective ones. An overview of the statistical characteristics and variability of key process variables is shown in Figure 4.

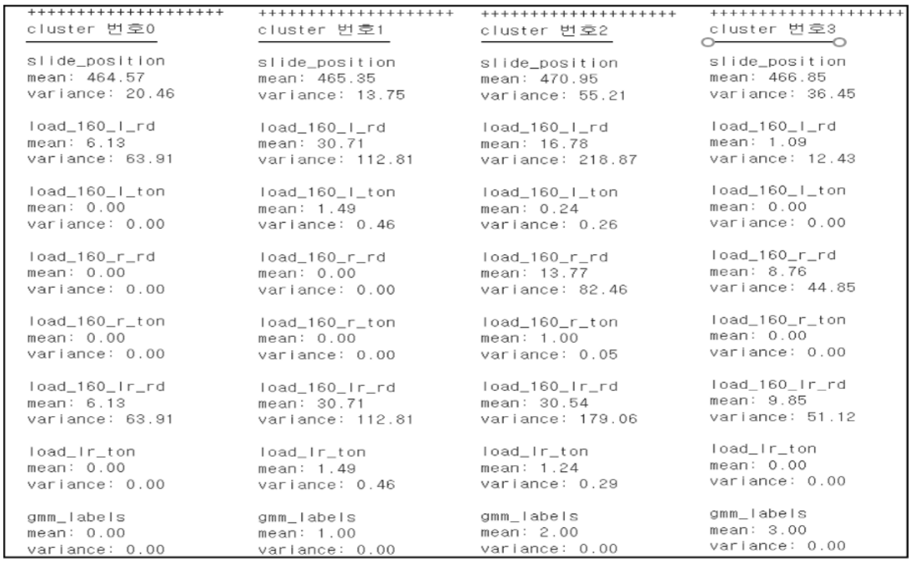


	slide_position	load_160_l_rd	load_160_l_ton	load_160_r_rd	load_160_r_ton	load_160_lr_rd	load_lr_ton
count	6719.000000	6719.000000	6719.000000	6719.000000	6719.000000	6719.000000	6719.000000
mean	466.590761	12.724215	0.428933	5.163864	0.168477	17.888079	0.597410
std	5.862247	15.104735	0.753854	7.656909	0.384905	14.723803	0.806161
min	439.030000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	462.310000	0.000000	0.000000	0.000000	0.000000	6.000000	0.000000
50%	463.120000	6.000000	0.000000	0.000000	0.000000	15.000000	0.000000
75%	469.090000	24.000000	1.000000	8.000000	0.000000	28.000000	1.000000
max	490.630000	67.000000	5.000000	38.000000	2.000000	76.000000	5.000000

Figure 4. Descriptive statistics of key injection molding process variables prior to clustering analysis.

The largest cluster contained the majority of data points and exhibited stable statistical characteristics across key process variables, indicating consistent and controlled operating conditions. This cluster was therefore identified as the normal process state. In contrast, the remaining three clusters showed distinct deviations in load- and motion-related variables, reflecting abnormal process behaviors associated with different defect patterns. These clusters captured variations that were not readily distinguishable using conventional threshold-based methods, highlighting the advantage of probabilistic clustering in complex manufacturing environments.

The soft assignment property of the GMM enabled a probabilistic interpretation of cluster membership, allowing gradual transitions between normal and abnormal states to be identified. This feature is particularly valuable for early-stage defect detection, where process conditions may begin to drift before fully developed defects occur. Overall, the clustering results demonstrate that the proposed GMM-based approach effectively separates normal and defective process states while maintaining stable and interpretable classification under nonlinear and non-normal data distributions. Figure 5 compares the statistical characteristics of process variables across different GMM clusters.



```
++++++++++++++++++++   ++++++++++++++++++++   ++++++++++++++++++++   ++++++++++++++++++++
cluster 번호0          cluster 번호1          cluster 번호2          cluster 번호3

slide_position         slide_position         slide_position         slide_position
mean: 464.57           mean: 465.35           mean: 470.95           mean: 466.85
variance: 20.46        variance: 13.75        variance: 55.21        variance: 36.45

load_160_l_rd          load_160_l_rd          load_160_l_rd          load_160_l_rd
mean: 6.13             mean: 30.71            mean: 16.78            mean: 1.09
variance: 63.91        variance: 112.81       variance: 218.87       variance: 12.43

load_160_l_ton         load_160_l_ton         load_160_l_ton         load_160_l_ton
mean: 0.00             mean: 1.49             mean: 0.24             mean: 0.00
variance: 0.00         variance: 0.46         variance: 0.26         variance: 0.00

load_160_r_rd          load_160_r_rd          load_160_r_rd          load_160_r_rd
mean: 0.00             mean: 0.00             mean: 13.77            mean: 8.76
variance: 0.00         variance: 0.00         variance: 82.46        variance: 44.85

load_160_r_ton         load_160_r_ton         load_160_r_ton         load_160_r_ton
mean: 0.00             mean: 0.00             mean: 1.00             mean: 0.00
variance: 0.00         variance: 0.00         variance: 0.05         variance: 0.00

load_160_lr_rd         load_160_lr_rd         load_160_lr_rd         load_160_lr_rd
mean: 6.13             mean: 30.71            mean: 30.54            mean: 9.85
variance: 63.91        variance: 112.81       variance: 179.06       variance: 51.12

load_lr_ton            load_lr_ton            load_lr_ton            load_lr_ton
mean: 0.00             mean: 1.49             mean: 1.24             mean: 0.00
variance: 0.00         variance: 0.46         variance: 0.29         variance: 0.00

gmm_labels             gmm_labels             gmm_labels             gmm_labels
mean: 0.00             mean: 1.00             mean: 2.00             mean: 3.00
variance: 0.00         variance: 0.00         variance: 0.00         variance: 0.00
```

Figure 5. Comparison of process variable statistics for each GMM cluster, distinguishing normal and defect-related states.

3.2. Defect Pattern Analysis

To further investigate the characteristics of defective production states, a detailed defect pattern analysis was conducted based on the clustering results obtained from the GMM. The three abnormal clusters exhibited distinct distributions across key process variables, indicating that different defect types are associated with unique combinations of process conditions rather than isolated parameter deviations. This observation underscores the importance of multivariate analysis for defect identification in injection molding processes. The distributions of representative process variables are illustrated in Figure 6, revealing skewed and imbalanced characteristics. These results indicate that defect formation in injection molding processes is governed by complex interactions among multiple process variables rather than isolated abnormalities. Each defect-related cluster exhibits characteristic combinations of load- and motion-related parameters, reflecting distinct underlying process instability mechanisms. Such cluster-specific patterns provide valuable insights into potential root causes of defects, including material flow imbalance, mechanical misalignment, and transient disturbances during critical molding stages. Consequently, the defect pattern analysis enables more targeted and interpretable process diagnosis, supporting proactive quality control and process optimization.

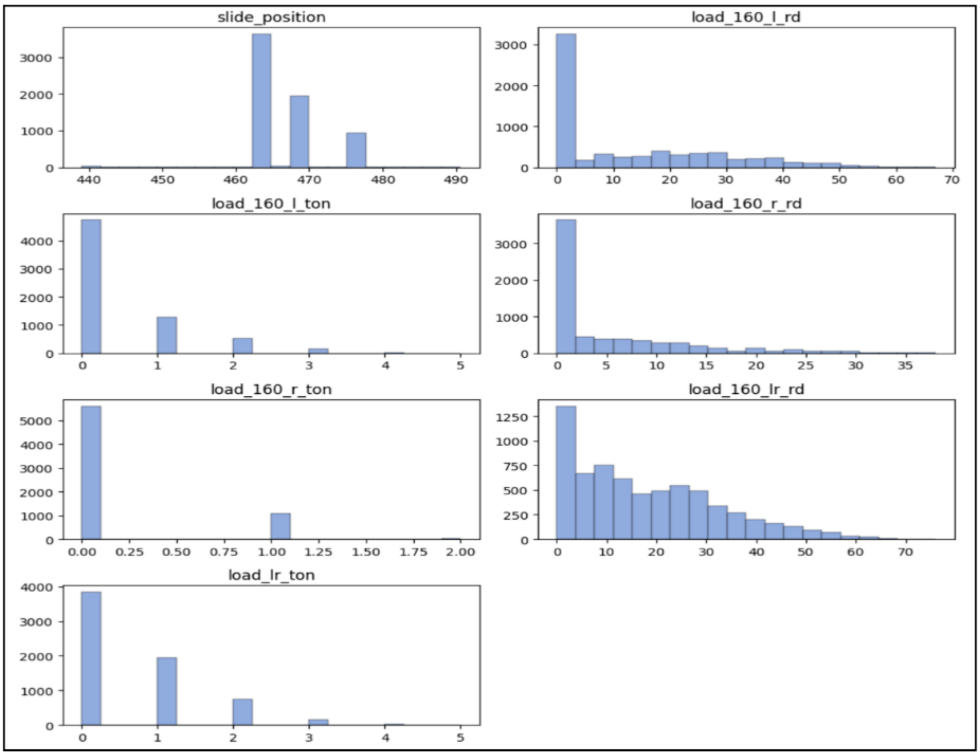
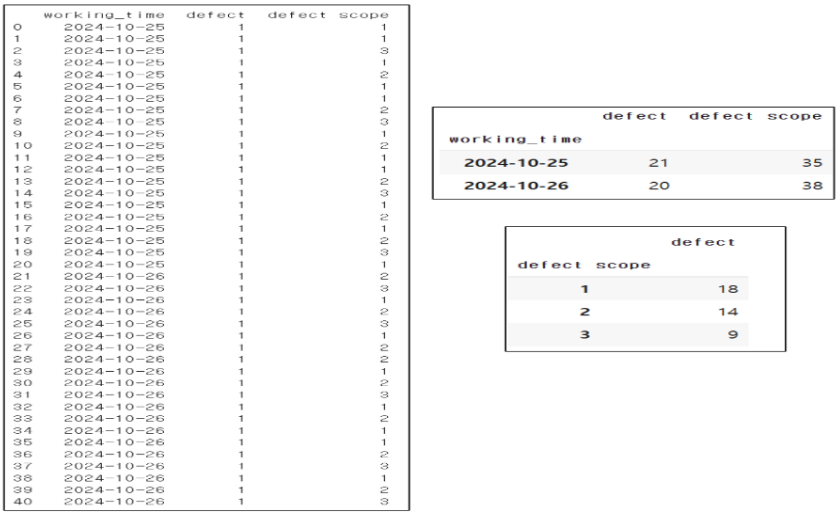


Figure 6. Histograms of selected process variables, showing skewed distributions and data imbalance in injection molding processes.

Cluster-level comparison revealed that variations in load-related variables were particularly prominent among defect clusters, reflecting unstable mechanical responses during critical stages of the molding cycle. Certain defect clusters were characterized by elevated load fluctuations, suggesting excessive pressure or material resistance, while others showed irregular motion-related patterns indicative of insufficient filling or improper mold closure. These differences imply that defect formation mechanisms may originate from diverse sources, such as material flow imbalance, mold misalignment, or transient disturbances in machine operation. Temporal patterns of defect occurrence across different defect clusters are shown in Figure 7.



```
    working_time  defect  defect scope
0     2024-10-25       1             1
1     2024-10-25       1             1
2     2024-10-25       1             3
3     2024-10-25       1             1
4     2024-10-25       1             2
5     2024-10-25       1             1
6     2024-10-25       1             1
7     2024-10-25       1             2
8     2024-10-25       1             3
9     2024-10-25       1             1
10    2024-10-25       1             2
11    2024-10-25       1             1
12    2024-10-25       1             1
13    2024-10-25       1             2
14    2024-10-25       1             3
15    2024-10-25       1             1
16    2024-10-25       1             2
17    2024-10-25       1             1
18    2024-10-25       1             2
19    2024-10-25       1             3
20    2024-10-25       1             1
21    2024-10-26       1             2
22    2024-10-26       1             3
23    2024-10-26       1             1
24    2024-10-26       1             2
25    2024-10-26       1             3
26    2024-10-26       1             1
27    2024-10-26       1             2
28    2024-10-26       1             2
29    2024-10-26       1             1
30    2024-10-26       1             2
31    2024-10-26       1             3
32    2024-10-26       1             1
33    2024-10-26       1             2
34    2024-10-26       1             1
35    2024-10-26       1             1
36    2024-10-26       1             2
37    2024-10-26       1             3
38    2024-10-26       1             1
39    2024-10-26       1             2
40    2024-10-26       1             3
```

```
              defect  defect scope
working_time
2024-10-25        21            35
2024-10-26        20            38
```

```
              defect
defect scope
1                 18
2                 14
3                  9
```

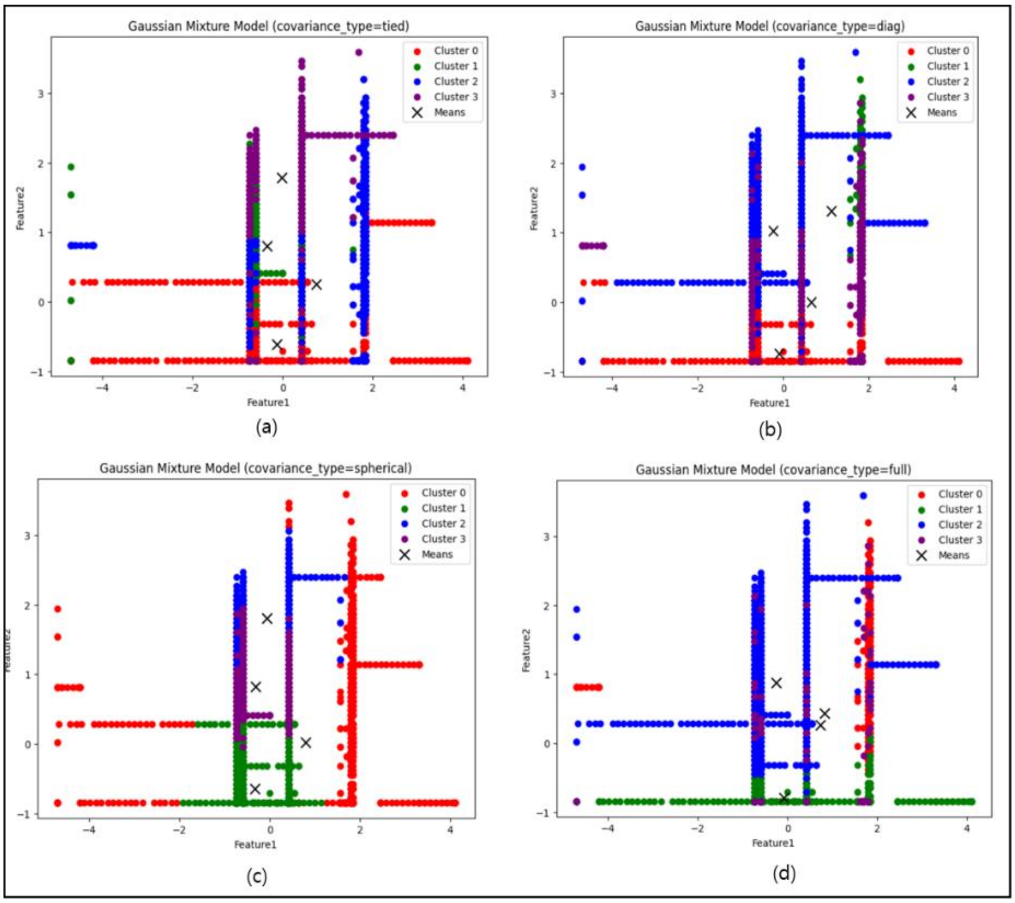
Figure 7. Temporal distribution of defect occurrences by defect cluster, indicating non-random defect patterns.

In addition, temporal analysis of defect cluster occurrence demonstrated that abnormal patterns were not randomly distributed but tended to emerge under specific operating conditions or process settings. This behavior indicates that defect patterns are closely linked to process dynamics and parameter interactions rather than stochastic noise. By capturing these nuanced variations, the proposed GMM-based framework enables a more comprehensive understanding of defect formation mechanisms and provides actionable insights for process adjustment and preventive quality control in injection molding environments.

3.3. Performance Evaluation and Discussion

The performance of the proposed GMM-based defect detection framework was evaluated in terms of stability, robustness, and practical applicability under real injection molding conditions. The clustering behavior remained consistent across repeated runs with different initialization conditions, indicating reliable convergence behavior. Unlike conventional threshold-based or fully supervised learning approaches, the proposed method does not rely on predefined defect labels, enabling consistent performance even in environments characterized by data imbalance and limited defect samples. The clustering results remained stable across repeated runs, indicating reliable convergence behavior and robustness to process variability.

From a manufacturing perspective, the proposed framework demonstrated a strong capability to capture subtle process deviations that precede visible defect formation. The probabilistic nature of the GMM allows gradual transitions between normal and abnormal states to be identified, supporting early-stage defect detection and preventive quality control. This characteristic is particularly beneficial for highly automated injection molding systems, where timely intervention can significantly reduce material waste, rework costs, and unplanned downtime. The clustering performance under different covariance structures is compared in Figure 8.



(a) (b)

(c) (d)

Figure 8. Comparison of GMM clustering results using different covariance structures: (a) tied, (b) diagonal, (c) spherical, and (d) full covariance.

Compared to traditional statistical quality control methods, which often struggle with nonlinear and multivariate data structures, the proposed approach provides a more flexible and adaptive mechanism for defect identification. Furthermore, in contrast to supervised machine learning models that require extensive labeled datasets, the unsupervised framework offers enhanced generalizability and scalability across different molding conditions and production lines. By reinforcing stable decision-making and preserving data integrity, the proposed method contributes to improved process reliability and manufacturing security.

Overall, the performance evaluation confirms that the proposed secure machine learning framework effectively supports defect reduction, process optimization, and operational efficiency in injection molding processes. These findings highlight its potential as a practical and reliable solution for data-driven quality control in smart manufacturing environments aligned with Industry 4.0.

4. Discussion

This study demonstrates that an unsupervised machine learning framework based on a Gaussian Mixture Model (GMM) can effectively support defect detection and quality enhancement in injection molding processes under realistic industrial conditions. Unlike conventional quality control approaches that rely on fixed thresholds or extensive labeled datasets, the proposed framework captures complex multivariate relationships among process variables and identifies defect-related patterns without requiring prior defect labeling. From a human-centric perspective, manufacturing security in this study is interpreted as the ability to provide reliable, trustworthy, and stable decision support for human operators under uncertain and variable production conditions. This characteristic is particularly advantageous in manufacturing environments where defective samples are scarce, imbalanced, or costly to annotate.

The clustering results highlight that defect formation in injection molding is not driven by isolated parameter anomalies but rather by combinations of process deviations involving load- and motion-related variables. By revealing distinct defect-related clusters with characteristic parameter patterns, the proposed approach provides deeper insight into underlying process dynamics compared to traditional statistical quality control methods. This pattern-based interpretation supports more informed decision-making and enables targeted process adjustments, which are essential for reducing defect rates and improving production stability.

From a performance and reliability perspective, the proposed GMM-based framework exhibits stable clustering behavior across repeated runs and varying initialization conditions. This stability indicates robust convergence and enhances confidence in the reproducibility of the defect detection results. In contrast to supervised machine learning models that may suffer from overfitting or performance degradation when operating conditions change, the unsupervised approach demonstrates greater adaptability to process variability, making it well suited for deployment in dynamic manufacturing environments.

In addition, the emphasis on data preprocessing, correlation analysis, and probabilistic clustering contributes to preserving data integrity and supporting secure decision-making in smart manufacturing systems. In this context, manufacturing security is interpreted as the ability to maintain reliable, stable, and trustworthy process monitoring and quality control under uncertain and evolving conditions. By reducing misclassification risks and supporting early-stage defect identification, the proposed framework aligns with the broader objectives of Industry 4.0, including resilient production systems and data-driven operational intelligence.

Despite these advantages, several limitations should be acknowledged. First, the number of clusters was determined based on domain knowledge and empirical observations, which may require adjustment for different products or molding conditions. Second, the analysis was conducted using data from a specific injection molding setup, and further validation across diverse machines and materials would enhance generalizability. Future research may extend the proposed framework by integrating adaptive cluster selection strategies, hybrid unsupervised–supervised learning schemes, or real-time implementation for online defect monitoring.

Overall, the discussion confirms that the proposed secure machine learning framework provides a practical, reliable, and scalable solution for defect detection and quality enhancement in injection molding processes, bridging the gap between data-driven analytics and industrial manufacturing requirements.

5. Conclusions

This study proposed a **secure machine learning framework** for defect detection and quality enhancement in injection molding processes based on an unsupervised Gaussian Mixture Model (GMM). By integrating systematic data preprocessing, correlation analysis, and probabilistic clustering, the proposed approach effectively identifies normal and defect-related process states without requiring labeled defect data. This characteristic addresses a critical limitation of conventional supervised learning methods and enhances applicability in real industrial environments characterized by data imbalance and limited defect samples.

The experimental results demonstrate that the proposed framework successfully captures complex multivariate relationships among process variables and distinguishes distinct defect patterns associated with abnormal operating conditions. The GMM-based clustering exhibited stable and consistent behavior across repeated runs, supporting reliable convergence and reproducible defect detection. Through pattern-based interpretation rather than single-parameter thresholding, the framework provides deeper insight into defect formation mechanisms and supports early-stage defect identification and preventive quality control.

From an industrial perspective, the proposed method contributes to improved process reliability, reduced defect occurrence, and enhanced operational efficiency in injection molding systems. Furthermore, by emphasizing data integrity, stable decision-making, and robustness to process variability, the framework aligns with the concept of **manufacturing security** in smart manufacturing environments. These features make the proposed approach well suited for deployment within Industry 4.0–oriented production systems.

Future work will focus on extending the framework to diverse molding conditions, materials, and product types, as well as incorporating adaptive cluster selection and real-time monitoring capabilities. Overall, this study demonstrates that secure and reliable unsupervised machine learning can play a pivotal role in advancing data-driven quality control and intelligent manufacturing systems.

Author Contributions: Conceptualization, M.-Y.K.; **methodology**, M.-Y.K.; **software**, M.-Y.K.; **validation**, M.-Y.K.; **formal** analysis, M.-Y.K.; **investigation**, M.-Y.K.; **resources**, M.-Y.K.; **data curation**, M.-Y.K.; **writing—original draft preparation**, M.-Y.K.; **writing—review and editing**, M.-Y.K.; **visualization**, M.-Y.K.; **supervision**, M.-Y.K.; **project administration**, M.-Y.K.; **funding acquisition**, M.-Y.K. The author has read and agreed to the published version of the manuscript.

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

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Folgado, F.J.; Calderón, D.; González, I.; Calderón, A.J. Review of Industry 4.0 from the Perspective of Automation and Supervision Systems: Definitions, Architectures and Recent Trends. *Electronics* **2024**, *13*, 782, doi:10.3390/electronics13040782.
2. Tao, F.; Qi, Q.; Liu, A.; Kusiak, A. Data-Driven Smart Manufacturing. *Journal of Manufacturing Systems* **2018**, *48*, 157–169, doi:10.1016/j.jmsy.2018.01.006.
3. Soori, M.; Arezoo, B.; Dastres, R. Internet of Things for Smart Factories in Industry 4.0, a Review. *Internet of Things and Cyber-Physical Systems* **2023**, *3*, 192–204, doi:10.1016/j.iotcps.2023.04.006.
4. Application of Machine Learning Techniques in Injection Molding Quality Prediction: Implications on Sustainable Manufacturing Industry Available online: https://www.mdpi.com/2071-1050/13/8/4120 (accessed on 26 December 2025).
5. Obregon, J.; Hong, J.; Jung, J.-Y. Rule-Based Explanations Based on Ensemble Machine Learning for Detecting Sink Mark Defects in the Injection Molding Process. *Journal of Manufacturing Systems* **2021**, *60*, 392–405, doi:10.1016/j.jmsy.2021.07.001.
6. Comparison of Hybrid Machine Learning Approaches for Surrogate Modeling Part Shrinkage in Injection Molding Available online: https://www.mdpi.com/2073-4360/16/17/2465 (accessed on 26 December 2025).
7. Zhang, Z.; Zhou, M.; Wan, H.; Li, M.; Li, G.; Han, D. IDD-Net: Industrial Defect Detection Method Based on Deep-Learning. *Engineering Applications of Artificial Intelligence* **2023**, *123*, 106390, doi:10.1016/j.engappai.2023.106390.
8. A Review on Industrial Surface Defect Detection Based on Deep Learning Technology | Proceedings of the 2020 3rd International Conference on Machine Learning and Machine Intelligence Available online: https://dl.acm.org/doi/10.1145/3426826.3426832 (accessed on 26 December 2025).
9. 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. *Mechanical Systems and Signal Processing* **2020**, *138*, 106587, doi:10.1016/j.ymssp.2019.106587.
10. Oluwasegun, A.; Jung, J.-C. A Multivariate Gaussian Mixture Model for Anomaly Detection in Transient Current Signature of Control Element Drive Mechanism. *Nuclear Engineering and Design* **2023**, *402*, 112098, doi:10.1016/j.nucengdes.2022.112098.
11. Advanced Anomaly Detection in Manufacturing Processes: Leveraging Feature Value Analysis for Normalizing Anomalous Data Available online: https://www.mdpi.com/2079-9292/13/7/1384 (accessed on 26 December 2025).

12. Sebestyen, H.; Popescu, D.E.; Zmaranda, R.D. A Literature Review on Security in the Internet of Things: Identifying and Analysing Critical Categories. *Computers* **2025**, *14*, 61, doi:10.3390/computers14020061.
13. Gyamfi, E.; Jurcut, A. Intrusion Detection in Internet of Things Systems: A Review on Design Approaches Leveraging Multi-Access Edge Computing, Machine Learning, and Datasets. *Sensors* **2022**, *22*, 3744, doi:10.3390/s22103744.
14. Umer, M.A.; Junejo, K.N.; Jilani, M.T.; Mathur, A.P. Machine Learning for Intrusion Detection in Industrial Control Systems: Applications, Challenges, and Recommendations. *International Journal of Critical Infrastructure Protection* **2022**, *38*, 100516, doi:10.1016/j.ijcip.2022.100516.
15. Liao, H.; Murah, M.Z.; Hasan, M.K.; Aman, A.H.M.; Fang, J.; Hu, X.; Khan, A.U.R. A Survey of Deep Learning Technologies for Intrusion Detection in Internet of Things. *IEEE Access* **2024**, *12*, 4745–4761, doi:10.1109/ACCESS.2023.3349287.
16. Cheng, Y.; Chen, W.; Fan, W.; Huang, W.; Yu, G.; Liu, W. IoTFuzzBench: A Pragmatic Benchmarking Framework for Evaluating IoT Black-Box Protocol Fuzzers. *Electronics* **2023**, *12*, 3010, doi:10.3390/electronics12143010.
17. Li, C.; Li, F.; Zhang, L.; Yang, A.; Hu, Z.; He, M. Intrusion Detection for Industrial Control Systems Based on Improved Contrastive Learning SimCLR. *Applied Sciences* **2023**, *13*, 9227, doi:10.3390/app13169227.
18. Gadallah, W.G.; Ibrahim, H.M.; Omar, N.M. A Deep Learning Technique to Detect Distributed Denial of Service Attacks in Software-Defined Networks. *Computers & Security* **2024**, *137*, 103588, doi:10.1016/j.cose.2023.103588.