EAI Endorsed Transactions
on Scalable Information Systems
Research Article **EAI.EU**

Deep Reinforcement Learning-Driven Adaptive Control for Intelligent Manufacturing Systems: A Multi-Sensor Fusion Framework for Real-Time Anomaly Detection

Shanshan Kong[1] Peng Zhou[2*]

[1,2]Shandong Huayu University of Technology, Dezhou, China, 253034

Abstract

One of the principal requirements for Intelligent Manufacturing Systems that are working in complicated, nonlinear, uncertain environments is the robust real-time monitoring and an adaptive control system. In this paper, a deep learning-based adaptive control framework is proposed that combines multi-sensor fusion, residual-based anomaly and robust state estimation. The framework is tested on a multi-sensor manufacturing dataset that consists of a total of time steps of four different heterogeneous sensor streams collected under both normal and anomalous operating The sensor data is initially preprocessed by noise filtering, normalization, synchronization, and feature extraction, after an Unscented Kalman Filter (UKF) is used for sensor fusion and state estimation. Residual analysis along with Vector Machine (SVM) allows for real-time anomaly detection and classification, while Multi-Model Adaptive Estimation (MMAE) technique improves the robustness of the state estimation. A Proximal Policy Optimization (PPO) agent uses updated system state and the anomaly information for adaptive control. The results of the experiments reveal perfect classification with an ROC-AUC of 1.000, highly stable state estimation in all state dimensions, and continuous performance increase with the average reward of 1.79, thereby validating the proposed framework's suitability for aware adaptive control in intelligent manufacturing systems.

Keywords: Deep Reinforcement Learning; Intelligent Manufacturing Systems; Multi-Sensor Fusion; Residual-Based Anomaly Detection; Adaptive Control; Proximal Policy Optimization

Received on 31 January 2026, accepted on 20 April 2026, published on 29 April 2026

doi: 10.4108/eetsis.11781

*Corresponding author. Email: PengZhou2026@outlook.com, zhoupeng19880923@163.com

1. Introduction

Industry 4.0 has been developing very fast, and one of its most remarkable effects is the conversion of the conventional manufacturing systems into smart, interconnected, and data-oriented environments [1]. The modern intelligent manufacturing system is the result of the combination of the cutting-edge technologies for sensing, automation, and AI, which together not only enhance the productivity but also the product quality and operational efficiency [2]. These systems are designed in such a way as to be constantly producing a vast amount of data that is diverse in nature through a wide- array of sensors keeping an eye on machine conditions, process variables, and environmental factors [3]. And while the data-rich settings make it possible for the most intelligent decision-making to be done, the complexity, uncertainty, and the necessity for real-time responsiveness are some of the challenges that come with such conditions [4]. Thus, guaranteeing the trustworthy operation through the prompt detection of anomalies and the responsive control has become a requirement that is of utmost importance in the industry and also in the domain of research [5].

Intelligent manufacturing and control paradigms that adapt were born to the limitations of conventional, rule-based and model-driven control strategies [6]. Traditional control methods are based on predefined mathematical models and fixed thresholds, which are often inadequate for capturing the complex, time-varying, and stochastic characters of modern manufacturing processes [7]. Besides, manufacturing systems contain a variety of sensors, which can measure vibration, temperature, current, pressure, and so on, and these sensors are heterogeneous ones with different sampling rates, noise levels, and quality of the signal they produce [8]. Moreover, the system faces challenges in monitoring and control due to sensor faults, data inconsistencies, and unexpected disturbances [9]. The said difficulties have led to the introduction of data-driven methods which can learn from operational data, merge information from different sensors, and change their behavior dynamically in response to environmental changes—all these without needing specific system models [10].

Real-time anomaly detection and adaptive control are two main pillars of modern-day intelligent manufacturing systems [11]. The absence of timely detection of anomalies or failures might result in a series of very undesirable consequences such as machine deterioration, involuntary stoppage of production, safety risks, and large financial losses[12]. Anomalous behavior gets quickly recognized through precise detection thus, allowing the use of preventive maintenance as well as fault remediation [13]. Meanwhile, adaptive control is a process that helps to keep the system performing at its best even when the operating conditions vary [14]. Robust state estimation in the presence of uncertainty and noise is what multi-sensor fusion techniques depend on, and thus, they are among the main contributors of enhancing system perception [15]. Deep reinforcement learning, on the other hand, has increasingly been recognized as empowering the decision-making and controlling process in an effective manner since agents get to learn the best policies through constant interaction with a complicated surrounding. Therefore, the integration of sensor fusion and reinforcement learning is viewed as a promising future pathway towards the realization of production systems that are fully automated, resilient, and efficient.

Even though the technologies have got better lately, the prevailing methods still treat anomaly detection, sensor fusion, and control optimization separately thus coming up with disconnected solutions that do not have comprehensive adaptability. A number of the anomaly detection methodologies rely wholly on supervised classification, which is disallowing the use of untagged data and having a hard time to apply the learning on unseen defect cases. Equally, the old control tactics may not be able to take full advantage of the real-time anomaly information in order to adjust the control actions in a preventive manner. In order to break through these limitations, the paper puts forward a single deep reinforcement learning–based initiative that combines adaptive control with multi-sensor fusion for real-time anomaly detection within the smart manufacturing systems. With the incorporation of sensor fusion and state estimation techniques, the proposed strategy is able to present reliable system representations, then amplifying the awareness of anomalies through residual-based methods and classification, and finally, having a deep reinforcement learning agent to dynamically optimize the control actions. By bringing perception, detection, classification, and decision-making together in a closed-loop manner, the proposed approach can not only face issues like uncertainty, nonlinearity, and real-time adaptability but also enhance the robustness, efficiency, and operational continuity in smart manufacturing environments. In addition, while deep reinforcement learning enables adaptive control in dynamic environments, ensuring closed-loop stability remains a critical challenge in nonlinear industrial systems. Future enhancements of the proposed framework may incorporate Lyapunov-constrained policy optimization techniques to provide theoretical stability guarantees and improve the reliability of reinforcement-driven control strategies.

1.1 Problem Statement

The detection and control of anomalies under complex industrial conditions in real-time, although intelligent manufacturing has progressed remarkably, are still the issues discussed in today's research Shahin et al.[16]. An earlier study has pointed out that the traditional model-based and rule-based methods of control have a hard time with adaptation to changes in such factors as nonlinear dynamics, sensor uncertainty, and disturbances Lian et al.[17]. Moreover, detection of anomalies in data has led to proposals of various methods, but many rely on high-quality labeled datasets and exhibit poor performance on fault scenarios that they have not seen before Bhanage et al.[18]. Moreover, the control mechanisms are often disregarded when developing sensor fusion systems which leads to unsynchronized systems that miss making timely decisions using anomaly data Desikan et al. [19].Deep reinforcement learning, despite being in its experimental phase for industrial control applications, has not yet gained full recognition as far as optimization is concerned, and for this reason, the combination of multi-sensor fusion and anomaly awareness has not been contemplated at the same time.

1.2 Objectives

The primary objective of this study is to develop an intelligent manufacturing framework that integrates multi-sensor fusion, anomaly detection, and deep reinforcement learning-based adaptive control.

Specifically, the study aims to:

- Design a robust multi-sensor fusion and state estimation framework using Unscented Kalman Filter (UKF) and Multi-Model Adaptive Estimation (MMAE) to improve system reliability under uncertainty.
- Develop a residual-based anomaly detection approach combined with Support Vector Machine (SVM) for accurate identification and classification of anomalous conditions.
- Construct an effective state representation by integrating fused sensor data, residual information, and anomaly classification outputs for control decision-making.
- Implement a Proximal Policy Optimization (PPO)-based deep reinforcement learning controller for adaptive and real-time control.
- Evaluate the proposed framework in terms of anomaly detection performance, state estimation accuracy, and control effectiveness.

1.3 Research Organization

The paper is organized into different sections as described below. Section 2 is dedicated to the literature review covering the existing studies in terms of multi-sensor fusion, anomaly detection, and deep learning-based control in intelligent manufacturing systems through with the help of the literature. In Section 3, the method is presented, which describes data cleansing, multi-sensor fusion along with UKF and MMAE, residual-based anomaly detection, and DRL-based adaptive control with PPO. The experimental setting is covered in Section 4 followed by the results from performance evaluation which consist of anomalies detection accuracy, the behavior of state estimation, and control effectiveness. Finally, Section 5 concludes the paper and discusses possible future research paths.

2. Literature Survey

In the study by Xiao et al.[20], a formal robot motion risk reasoning framework was suggested to overcome the limitations of currently available risk-aware path planning techniques. Present methods that usually depend on arbitrary risk functions or chance constraints have been considered inadequate mainly because of their low accuracy and excessive caution. The authors present a detailed risk paradigm that includes local, action, and traverse-dependent risks which makes it possible to create a risk-aware planner. The experimental results confirm that the framework can find safer and more accurate paths in unstructured environments.

The Adaptive Adversarial Transformer is introduced by Orabi et al.[21] as a novel method for anomaly detection in smart manufacturing from the Industry 5.0 perspective. Traditional methods struggle with high-dimensional, non-stationary time-series data and therefore often overlook subtle anomalies. It is a state-of-the-art model that integrates Transformer architecture, anomaly attention mechanisms, and Adversarial Learning to discover intricate temporal patterns and change thresholds instantaneously. The validation performed with both the ordinary and real-life datasets produced F1 scores that were above those of the other models and this is a strong indication of Its outstanding performance.

In their research work, Wong et al.[22] proposed a multi-sensor fusion scheme that included deep reinforcement learning (DRL) and multi-model adaptive estimation (MMAE) as a method to enhance the simultaneous localization and mapping (SLAM) process. One of the main tasks is the integration of LiDAR-based PLICP with RGB-D camera-based ORBSLAM2 through Proximal Policy Optimization (PPO), which allows the optimal sensor weight modification. The results from the test runs reveal that the mobile robot's localization accuracy and stability in difficult environments have been significantly improved, thus giving the robots the upper hand over the traditional SLAM methods.

According to Zhou et al.[23] propose a method called reinforcement-learning-based-data-fusion (RLBDF) which is specifically designed to enhance the reliability and strength of multi-sensor systems during air combat. The method employs cubic B-spline interpolation to solve time alignment issues and increases the precision of fusion by applying an error correction technique between the fused and true values. In the absence of prior knowledge, Fisher information is used, in turn, as a reward. Simulation results verify the technique's efficiency and practicality in the case of air combat multi-sensor data fusion.

Zhe Sun et al . [24] introduce an autonomous UAV obstacle avoidance technique that merges multi-sensor fusion and deep reinforcement learning for path optimization. The proposed approach combines the enhanced Bayesian fusion method for coordinated multi-UAV inspection and SE-f-AnoGAN for the detection of the faulty power inspection images. The attention mechanisms applied to the anomaly detection model result in better accuracy and precision of the system. The DQN-driven route planning guarantees no collision, and quick inspections with very good output results for different datasets.

In this article, Yu and Liu [25] introduced algorithms for sensor network direction control using deep reinforcement learning (DRL).Their proposal, in contrast to standard techniques which require set target models, gives the freedom to the sensors to regulate their actions by themselves, thus getting the balance right between tracking and searching. The use of the multiagent deep deterministic policy gradient algorithm provides stability in decentralized decision-making. The technique, in conjunction with a labeled multi-Bernoulli filter for multitarget tracking proves to be very efficient during difficult simulation tests.

Alginahi et al. [26] offered a bibliometric review that depicted the advancements in Reinforcement Learning and Deep Reinforcement Learning for Industrial Automation purposes. They analyzed 672 articles from peer-reviewed journals for the years 2017 to 2026 and discovered major trends such as the 42% portion of DRL and 27% portion of Multi-Agent RL in robotic control and process optimization areas, respectively. Even though the review has shown a significant growth potential, it has also mentioned the issues of lack of real-world implementation, limited scalability, and safety validation as the major points in the path of RL to the industry.

A pioneering method for handling concurrency in Federated Digital Twins (FDTs) was presented by Anwar et al. [27] as the one that would be directly applicable to the current manufacturing industry. The proposed technique is based on the combination of Deep Reinforcement Learning (DRL) and the Priority Ceiling Protocol (PCP) to deal with issues such as priority inversions and failures in synchronization that are typical of situations that require real-time and dynamic responses. The technique has been evaluated through SimPy-based discrete-event simulations that not only improve the distribution of resources but also result in up to 24.27% reduction in delivery times and 6.65% decrease in waiting time for urgent jobs as opposed to the traditional PCP controllers, thus, drastically increasing the overall efficiency of the production process.

The Huang et al. [28] have introduced a new method for reconfiguration planning which is based on Deep Reinforcement Learning for reconfigurable machine tools (RMTs) digital twin-assisted smart manufacturing systems. The method resolves the problems of dynamic reconfiguration by reducing the expenses of reconfiguration, moving, and processing to the minimum. For exploring the state and action spaces aimed at discovering the optimal policy, the Deep Q-Network (DQN) is employed. A case study in the industry showcases the capability of the method to reconfigure RMTs and thus to produce several parts efficiently in smart manufacturing settings.

Most of the research works done so far have dealt with anomaly detection, sensor fusion, or control optimization separately, which has resulted in a very limited applicability of these techniques to real-time manufacturing systems. Very infrequently does malfunction information come to be a part of the control decisions and the deep reinforcement learning (DRL) methods usually take error-free state estimates as their point of departure. In order to tackle these drawbacks, the present study puts forward a comprehensive system that interlinks multi-sensor fusion, residual-based anomaly detection, and DRL-based adaptive control within a closed-loop system.

3. Methodology

The proposed framework is a combination of multi-sensor fusion, residual-based anomaly detection, and deep reinforcement learning for adaptive control in smart manufacturing systems shown in Figure 1. The data from all the sensors is pre-processed before it is combined, consisting of filtering, normalization, synchronization, and feature extraction, and then it is combined using an Unscented Kalman Filter (UKF) and even further augmented by Multi-Model Adaptive Estimation (MMAE). Detecting anomalies occurs through the computation of residuals, and the latter is categorized with the assistance of Support Vector Machine (SVM). The real-time Proximal Policy Optimization (PPO) agent is the one in charge of optimization. The system operates in such a way that it is resilient, adaptable, and capable of withstanding anomalies, and this is achieved by taking the polished state, and residuals, labeling anomalies, and recording past control actions as inputs.

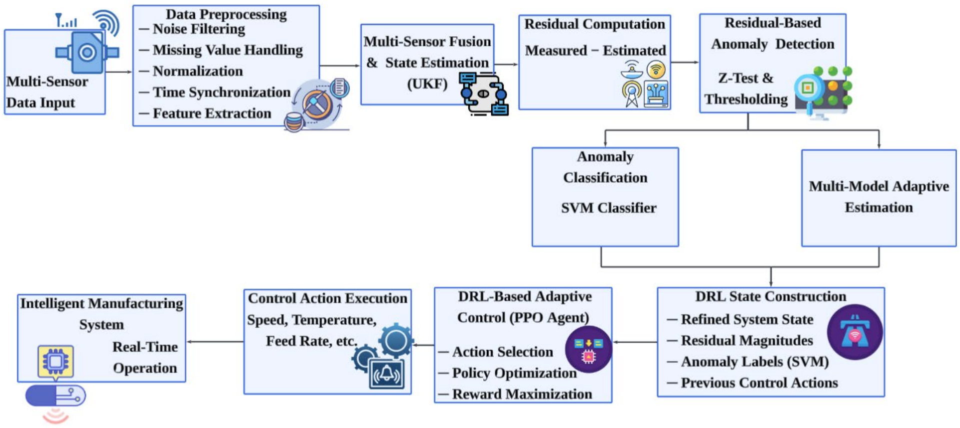

Figure 1. Workflow Diagram

3.1 Data Collection

The collected data that were available online, such as the industrial pump monitoring dataset at Kaggle provided multi-sensor data in this investigation. It is a collection of multivariate time-series sensor signals recorded from the industrial pumping system operating under both normal and faulty conditions. Specifically, it consists of various sensor measurements like temperature, vibration, current (power), and pressure/flow which are recorded for 10,000 time steps. The data from the different sensor sources are first combined into a unified multivariate time-series format before sensor fusion is carried out. This integration process deals with the diversity of signal characteristics, sampling behavior, and noise levels, thus ensuring temporal alignment and consistency among all sensor streams. This pre-integration is very crucial for the effective multi-sensor fusion, robust state estimation, and the following anomaly-aware adaptive control.

https://www.kaggle.com/datasets/nphantawee/pump-sensor-data

Data preprocessing is an essential step in intelligent manufacturing systems. Raw sensor data are usually collected from industrial environments that are very noisy, incomplete, and heterogeneous. The harsh operating conditions and sensor imperfections are the main reasons for such data quality. In this study, we perform preprocessing to improve data quality and support reliable analysis downstream. Noise filtering methods are used to suppress high-frequency noise and disturbances which are not part of the main signal while maintaining its essential features.

Missing data coming from sensor faults or communication delays are managed with proper interpolation or imputation techniques to keep the data continuous. Normalization is applied to make the scale of different sensor measurements equal, thus, the standardization of feature ranges will prevent bias in learning algorithms. Time synchronization is the aligning of multi-sensor data streams with different sampling rates into a common temporal framework, thus enabling accurate sensor fusion. Afterwards, feature extraction takes place and from raw signals, informative statistical and temporal features, such as mean, variance, and frequency-domain attributes, are extracted thereby reducing dimensionality and improving the process of anomaly detection, classification, and adaptive control which benefits the latter because of the reduction in dimensionality.

3.1.2 Noise filtering

Noise filtering is a fundamental operation that precedes processing and its main goal is to get rid of unwanted interferences and measurement noise from raw sensor signals but without losing the underlying system dynamics. In the smart factories, sensor data are often contaminated with high-frequency noise, which can be attributed to mechanical vibrations, electrical interference, and environmental changes. To get rid of these noise signals, a digital low-pass filter is used to attenuate the high-frequency components that are outside the operational bandwidth of the production process.

The filtered signal $\tilde{x}(t)$ is obtained as shown in Eqn. (1).

$$\tilde{x}(t) = x(t) * h(t) \quad (1)$$

Here $x(t)$ represents the raw sensor signal, $h(t)$ denotes the filter impulse response, and $*$ indicates convolution. In the frequency domain, the filtering operation can be expressed as shown in Eqn. (2).

$$\tilde{X}(f) = X(f) \cdot H(f) \quad (2)$$

Here $X(f)$ and $H(f)$ are the Fourier transforms of the signal and filter, respectively. Alternatively, a moving average filter may be used for real-time applications, defined as shown in Eqn. (3).

$$\tilde{x}_k = \frac{1}{N}\sum_{i=0}^{N-1} x_{k-i} \quad (3)$$

Here N is the window size. These filtering techniqu, effectively enhance signal quality and improve the reliability of subsequent sensor fusion, anomaly detection, and control processes.

The selection of low-pass and moving average filters is motivated by their effectiveness in preserving essential signal characteristics while suppressing high-frequency noise commonly observed in industrial sensor data. The cutoff frequency of the low-pass filter is chosen based on the operational bandwidth of the manufacturing process to ensure that relevant system dynamics are retained. Similarly, the window size of the moving average filter is selected to balance noise reduction and signal responsiveness, avoiding excessive smoothing that may distort transient behaviors. Compared to more complex filtering techniques, these methods offer computational efficiency and real-time applicability, making them suitable for heterogeneous multi-sensor environments with varying noise profiles.

3.1.3 Missing Value Handling

Handling missing values is an important preprocessing step in AI-powered manufacturing systems. The data from sensors can have interruptions caused by communication problems, faulty sensors, or errors in data collection. If missing values are not treated, they will cause distortion in statistical properties and deteriorate the quality of sensor fusion, anomaly detection, and learning algorithms. The authors of the current research handle the missing values by applying interpolation and imputation techniques to keep the data continuous. For data gaps that last a short period, linear interpolation is utilized, which is defined as shown in Eqn. (4).

$$\tilde{x}(t) = x(t_1) + \frac{x(t_2)-x(t_1)}{t_2-t_1}(t-t_1), t_1 < t < t_2 \quad (4)$$

Here $x(t_1)$ and $x(t_2)$ are the nearest observed samples before and after the missing interval. For longer gaps, mean imputation may be used, expressed as shown in Eqn. (5).

$$\tilde{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (5)$$

Here N denotes the number of available samples. These methods ensure data completeness while minimizing distortion, thereby enabling robust downstream processing and real-time system analysis.

The choice of linear interpolation and mean imputation is based on the duration and characteristics of missing data segments. Linear interpolation is applied for short-duration gaps to preserve temporal continuity and maintain local trends in sensor signals, thereby minimizing distortion in statistical properties. In contrast, mean imputation is used for longer gaps where interpolation may introduce unrealistic patterns, ensuring stability in the overall data distribution. These methods help maintain statistical consistency of the dataset and prevent bias in feature extraction. Furthermore, by reducing abrupt discontinuities and preserving signal characteristics, the adopted approach supports reliable residual computation and improves the robustness of downstream anomaly detection and classification processes.

3.1.4 Normalization

Normalization takes place to bring the varying sensor readings to a common numerical range, thereby preventing any one feature from completely taking over the learning and estimation processes owing to its higher magnitude or different units. In smart manufacturing systems, sensor measurements like temperature, vibration, current, and pressure usually have very different scales which can negatively impact the application of sensor fusion, classification, and reinforcement learning methods. The authors of this paper choose min-max normalization because of its straight-forwardness and great performance in real-time use. The formula for the normalized signal is given as where x symbolizes the original sensor value, x min and x max depict the minimal and maximal values of the corresponding feature, respectively. With this transformation, all features are rescaled to the range [0,1], which results in better numerical stability and faster model convergence. Min–max normalization fits very well with deep reinforcement learning and support vector machine classifiers, as it maintains the relative relationships among data points and at the same time allows for seamless multi-sensor data fusion.

3.1.5 Time synchronization

In multi-sensor intelligent manufacturing systems, time synchronization is very crucial because different types of sensors are usually working at various sampling rates and undergoing different communication delays. If synchronization is not performed, then mismatched data streams would result in imprecise sensor fusion and unreliable state estimation. For the purpose of this research, time synchronization is carried out by resampling all sensor signals to a common reference timeline with the help of interpolation.

Let $x_i(t_i)$ denote the measurements from the i-th sensor sampled at time instants t_i. The synchronized signal $\tilde{x}_i(t)$ at a unified time t is obtained using linear interpolation, expressed as shown in Eqn. (6).

$$\tilde{x}_i(t) = x_i(t_k) + \frac{x_i(t_{k+1}) - x_i(t_k)}{t_{k+1} - t_k}(t - t_k), t_k \leq t \leq t_{k+1} \tag{6}$$

Here t_k and t_{k+1} are the nearest sampling instants surrounding t. By this method, all sensor readings are synchronized to a single-time references, which facilitates flawless integration of sensors, computation of residues, and detection of anomalies and control adjustments in real-time.

3.1.6 Feature Extraction

Feature extraction is the process of converting raw time series data from multiple sensors into a compact and informative representation that embodies the vital features of the manufacturing system. The use of raw sensor signals directly results in high dimensionality and increased computational complexity, which can negatively affect the performance of sensor fusion, anomaly detection, classification, and reinforcement learning models. In this research, statistical time-domain features are calculated from each signal of the sensor throughout a moving time frame. Typical features are the mean, standard deviation, and root mean square (RMS), which are expressed by the following definitions respectively shown in Eqn. (7-9).

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{7}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2} \tag{8}$$

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{9}$$

Here x_i indicates the sensor readings that are taken within a window of size N. Also, frequency-domain characteristics like spectral energy can be computed with the help of the Fast Fourier Transform (FFT) which is intended to exhibit the dynamics of the system. Such characteristics obtained from extraction process are less affected by noise, have lower dimensions, and are more useful in multi-sensor fusion, anomaly detection, and adaptive control in the proposed framework compared to other techniques.

The handcrafted statistical features provide an efficient and interpretable representation; they may not fully capture complex nonlinear interactions among multi-sensor signals. To further enhance feature expressiveness, representation learning techniques such as autoencoder-based encoders or deep feature extraction models can be integrated prior to statistical feature computation. Such approaches enable the learning of compact and informative latent representations, thereby improving the modeling of cross-sensor dependencies and enhancing downstream tasks such as anomaly detection and adaptive control.

3.1.7 Data Splitting

After the data preprocessing stages, the multivariate time-series data are separated into training and testing datasets which will be used for the respective processes of model development and evaluation. In this case, a chronological split is employed rather than random sampling in order to preserve the temporal dependencies that are characteristically associated with industrial sensor data. In particular, 70% of the data are earmarked for the training of the sensor fusion, anomaly detection, and deep reinforcement learning models, while the remaining 30% are reserved for testing purposes and performance assessment. This approach eliminates the risk of information leaking and allows for a genuine evaluation under real-time operating conditions.

The chronological 70–30 data split is designed to preserve the temporal structure of the dataset while ensuring representation of different operational regimes across both training and testing sets. The dataset encompasses varying system conditions, including normal operation and fault scenarios, distributed over the 10,000-time steps. By maintaining temporal ordering, the model is evaluated under realistic conditions where future states are predicted based on past observations. This approach also helps capture potential regime-dependent variations and gradual system transitions, thereby providing a more reliable assessment of model generalization under evolving manufacturing dynamics.

3.2 Multi-Sensor Fusion and State Estimation

The multi-sensor fusion and state estimation process results in a reliable and comprehensive depiction of the manufacturing system by merging the information obtained from diverse sensors that operate with uncertainty. In a factory, each sensor output could be inaccurate, could cover only a small part of the whole situation, or might have very low trustworthiness; therefore, the incorporation of different sensor sources yields better accuracy and more reliable results. The study follows the nonlinear state estimation technique using an Unscented Kalman Filter (UKF). The system dynamics are described as

Kalman Filter (UKF) is adopted for nonlinear state estimation. The system dynamics are modeled as shown in Eqn. (10).

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{w}_{k-1} \tag{10}$$

and the measurement model is expressed as shown in Eqn. (11).

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k \tag{11}$$

Here $\mathbf{x}_k$ denotes the system state vector, $\mathbf{u}_{k-1}$ represents the control input, $\mathbf{z}_k$ is the fused sensor measurement vector, and $\mathbf{w}_{k-1}$ and $\mathbf{v}_k$ are process and measurement noise, respectively. The UKF propagates sigma points through the nonlinear functions $f(\cdot)$ and $h(\cdot)$ to estimate the posterior state mean and covariance. This fusion framework dismantles uncertainty efficiently, improves the precision of state estimation, and gives a strong support for the residual-based anomaly detection and adaptive control methods. The measurement noise covariance is assumed to be constant in the Unscented Kalman Filter. However, in real-world manufacturing environments, sensor reliability may vary over time due to noise, degradation, or environmental disturbances. To address this limitation, adaptive measurement covariance tuning strategies can be incorporated to dynamically adjust sensor weighting based on real-time credibility, thereby enhancing the robustness and accuracy of the multi-sensor fusion process.

3.3 Residual-Based Anomaly Detection

Residual-based anomaly detection is a model-driven approach that detects abnormal behavior of the system by studying the discrepancy between the sensor readings taken and their computed values. In smart manufacturing plants, a state estimator like the Unscented Kalman Filter predicts measurements according to the learned system dynamics. The residual signifies the disparity between the true sensor reading and the estimated measurement, mathematically represented as shown in Eqn. (12).

$$\mathbf{r}_k = \mathbf{z}_k - \hat{\mathbf{z}}_k \tag{12}$$

where z_k indicates the acquired sensor vector and $\hat{\mathbf{z}}_k$ indicates the projected measurement derived from the state estimator. In normal conditions, the residuals have a distribution that is very close to zero in mean and has a bounded variance. Statistical thresholding using the Z-test, which computes the standardized residual, is employed to detect anomalies. The selection of the Z-test threshold is guided by a trade-off between detection sensitivity and false-alarm rate. A lower threshold increases sensitivity to subtle anomalies but may lead to higher false positives, whereas a higher threshold reduces false alarms at the cost of missed detections. In practice, the cutoff value is determined based on statistical confidence intervals of the residual distribution under normal operation, typically aligned with standard deviation multiples. This choice ensures a balance between timely anomaly detection and operational reliability, while maintaining computational simplicity suitable for real-time industrial monitoring environments, where the standardized residual is computed as shown in Eqn. (13).

$$Z_k = \frac{r_k - \mu_r}{\sigma_r} \tag{13}$$

The mean μ_r and standard deviation σ_r are utilized to characterize the residuals in normal operation. When the absolute Z-score surpasses a certain limit, the system is deemed to be in a state of anomaly. This technique facilitates very rapid detection of the above, namely faults, sensor failures, and non-standard behaviors in the process, with only a minimal quantity of labeled data required. First, the anomalies are discovered through the Z-test-based thresholding technique applied to the residuals. Then, a classifier using Support Vector Machine (SVM) is used to assign the discovered anomalies to the classes of either normal or faulty, which allows more precise adaptive control decisions to be made.

The normal operating conditions, the residuals are generally assumed to follow a zero-mean Gaussian-like distribution due to the stochastic nature of process and measurement noise. This statistical characterization enables the use of Z-score-based thresholding, where thresholds are calibrated based on the estimated mean and variance of residuals under nominal conditions. However, variations in sensor noise levels can influence the spread of the residual distribution, thereby affecting detection sensitivity. To address this, threshold selection must account for noise-induced variance changes to avoid false alarms or missed detections. This understanding improves the robustness of anomaly detection under varying sensor noise conditions.

The Z-score-based statistical thresholding provides an effective mechanism for anomaly detection; it assumes stationary statistical properties of the residuals. In practical industrial environments, process dynamics may vary over time, leading to non-stationary behavior. To enhance robustness, adaptive statistical profiling techniques, such as time-varying mean and variance estimation, can be incorporated to dynamically adjust detection thresholds and improve anomaly detection performance under evolving operating conditions.

3.3.1 SVM-Based Anomaly Classification

Support Vector Machine (SVM) is an algorithm that performs classification by the residual-derived features of the detected anomalies show in Figure 2. It is a supervised learning model that finds the hyperplane with the best margin between the two classes of operating conditions, normal and anomalous, thus separating them. In the current research, the amplitude and statistical characteristics of the residual coupled with the output from the multi-sensor data form the basis of the input features for the SVM classifier. The result from the SVM classifier is a binary output that marks states as either normal or anomalous. EMU can run at a maximum speed of 192 Km/h and could brake to a halt from this speed within 30 seconds. The performs residual-based anomaly detection followed by SVM-based classification in a sequential manner; this separation may introduce intermediate inconsistencies between detection and labeling stages. To improve coherence, future enhancements may consider joint optimization strategies that tightly integrate residual inference with SVM decision boundaries. Such an approach can reduce inconsistency between intermediate decisions and enhance the continuity and reliability of anomaly-aware adaptive control.

The single dataset configuration, further evaluation under cross-condition generalization scenarios can provide a more comprehensive assessment of the model's robustness. Future work may include training the SVM model under specific operating conditions and testing it on unseen or varying fault scenarios to better demonstrate its discrimination capability across diverse industrial environments.

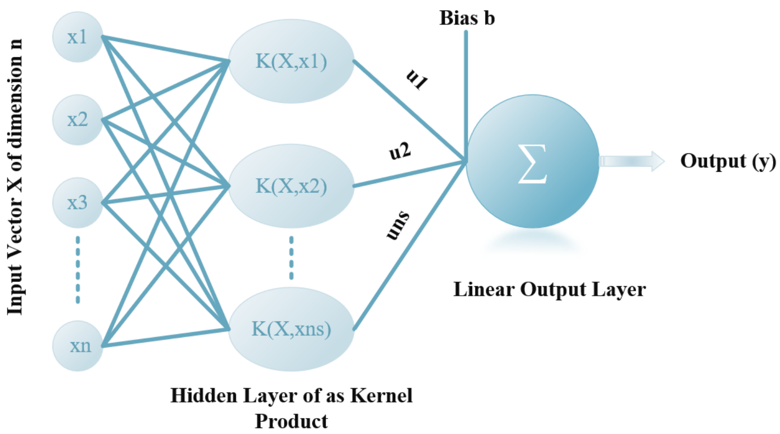

Figure 2. SVM Architecture Diagram

3.4 Robust State Refinement Using Multi-Model Adaptive Estimation

MMAE is used to obtain solid state refinement which improves the accuracy of estimation in the presence of changing operational conditions and possible sensor faults. In smart manufacturing systems, one system model might not be enough to cover different operating modes such as normal operation, degradation, or fault conditions. MMAE solves this problem by executing several state estimators concurrently, with each associating with a unique presumed system model. For the i th model, the state estimate $\hat{\mathbf{x}}_k^{(i)}$ and innovation covariance $\mathbf{S}_k^{(i)}$ are obtained using an Unscented Kalman Filter. The likelihood of each model is computed as shown in Eqn. (14).

$$\Lambda_k^{(i)} = \frac{1}{\sqrt{(2\pi)^n \left|\mathbf{S}_k^{(i)}\right|}} \exp\left(-\frac{1}{2}\mathbf{r}_k^{(i)\top}\mathbf{S}_k^{(i)-1}\mathbf{r}_k^{(i)}\right) \tag{14}$$

where $\mathbf{r}_k^{(i)}$ is the residual of the i-th model. The model probabilities are then updated as shown in Eqn. (15).

$$p_k^{(i)} = \frac{\Lambda_k^{(i)} p_{k-1}^{(i)}}{\sum_{j=1}^{M} \Lambda_k^{(j)} p_{k-1}^{(j)}} \tag{15}$$

Finally, the state of the refined system is obtained by a weighted fusion of individual model estimates shown in Eqn. (16).

$$\hat{\mathbf{x}}_k = \sum_{i=1}^{M} p_k^{(i)} \hat{\mathbf{x}}_k^{(i)} \tag{16}$$

The refinement based on MMAE strengthens the robustness of the system against uncertainties related to the model and faults in the sensors, thereby giving a trustworthy state representation for both anomaly detection and adaptive control.

The integration between the Unscented Kalman Filter and the MMAE framework follows a probabilistic model selection mechanism, where each candidate model is associated with an independent UKF estimator. At each time step, residuals generated by individual UKFs are used to compute likelihood functions, which in turn update the model probabilities. The model selection is implicitly governed by the maximum a posteriori probability, while maintaining a soft switching behavior through weighted state fusion. As system dynamics evolve, the probability distribution adapts continuously, enabling smooth transitions between models without abrupt switching. This probabilistic transition logic ensures robust tracking of changing operating conditions and enhances adaptability under dynamic manufacturing environments.

In the MMAE framework, each candidate model is designed to represent a distinct operating condition of the manufacturing system, such as normal operation, gradual degradation, and fault scenarios. These models differ in their system dynamics, noise characteristics, or parameter configurations, allowing them to capture diverse system behaviors. As the system transitions between operating modes, the associated model probabilities evolve accordingly, providing an interpretable indication of the underlying system condition. This structural differentiation enhances the understanding of mode transitions and supports more reliable decision-making in anomaly-aware adaptive control.

The current state representation incorporates the estimated system state, residuals, anomaly labels, and past actions, it does not explicitly account for estimation uncertainty. To further enhance decision-making under stochastic disturbances, future extensions may integrate covariance-aware belief information derived from the Unscented Kalman Filter into the policy input. This would enable the reinforcement learning agent to perform uncertainty-sensitive adaptive control by considering both state estimates and their associated confidence levels.

3.5 DRL State Construction

DRL state construction specifies the way that important system data is illustrated and given to the reinforcement learning agent for making decisions. In the case of intelligent manufacturing systems, a good state representation needs to include the present operating condition, anomaly data, and historical control context. This research develops the DRL state vector by merging the polished system state from multi-model adaptive estimation, residual-based anomaly indicators, classification results, and previous control actions. The state at the time step t is characterized as shown in Eqn. (17).

$$\mathbf{s}_t = [\hat{\mathbf{x}}_t^{\mathrm{MMAE}}, \mathbf{r}_t, y_t, \mathbf{a}_{t-1}] \tag{17}$$

Here $\hat{\mathbf{x}}_t^{\mathrm{MMAE}}$ is the robustly estimated system state, is the residual vector,
y_t is the anomaly classification label from SVMand $\mathbf{a}_{t-1}$ is the previous control action. This kind of state representa tion guarantees that the DRL agent is fully aware of the sit uation, which help it to learn the control policies that are a dapted to the system dynamics, anomaly severity, and past actions. Thus, the created state space support stable learni ng and efficient real-time control optimization. The formulation primarily relies on instantaneous features, which may limit the ability of the agent to capture gradual system drifts and temporal dependencies. To address this limitation, temporally enriched representations, such as sliding window-based state augmentation or sequence-aware feature encoding, can be incorporated to include historical context. This enhancement enables the reinforcement learning agent to better capture progressive

drift phenomena and improve adaptive control performance under evolving manufacturing conditions.

Here, a deep reinforcement learning agent based on Proximal Policy Optimization (PPO) is created to execute adaptive control by means of the developed state representation. The PPO agent gets as input the very reliably estimated system state, the magnitudes of the residuals showing how serious the anomaly is, the anomaly category that has been labelled by the SVM, and the previous control actions. The agent, utilizing such an extensive state information, delivers continuous control actions that are adapting in real time to the changing operating conditions and the detected anomalies. The PPO framework provides stable policy updates through the use of clipped objective optimization which in turn makes control adaptation in complicated manufacturing environments to be both reliable and efficient. The choice of the Proximal Policy Optimization (PPO) algorithm in this research is motivated by its reliability and not by its strength in control tasks in continuous operation shown in Figure 3. The DRL agent monitors the manufacturing environment by perceiving the present state s_t, choosing a control action a_t , and getting a reward r_t that indicates how well the system is performing. The agent's goal is to get the highest expected cumulative discounted reward, which is specified as shown in Eqn. (18).

$$J(\theta) = \mathbb{E}[\sum_{t=0}^{T} \gamma^t r_t] \tag{18}$$

Here $\gamma \in (0,1)$ is the discount factor and θ represents the policy parameters. PPO updates the policy by optimizing a clipped surrogate objective function, expressed as shown in Eqn. (19).

$$L^{\text{PPO}}(\theta) = \mathbb{E}\left[\min\left(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t\right)\right] \tag{19}$$

Here $r_t(\theta)$ is the probability ratio between new and old policies, $\hat{A}_t$ is the advantage estimate, and ϵ is the clipping parameter. The DRL agent gets to know adaptive control methods that manage disturbances, improve the process stability, and incrementally upgrade the manufacturing productivity in real time through back-and-forth engagement and policy improvement. The primarily focuses on system performance; it does not explicitly incorporate anomaly severity. To enhance safety-oriented control, future improvements may consider integrating residual magnitude into the reward structure, where higher residual values corresponding to abnormal conditions impose penalties on the agent. This anomaly-sensitive reward design can encourage safer control actions and improve robustness during fault-prone operating intervals.

The computational complexity of the proposed framework is influenced by the integration of multiple modules, including UKF-based state estimation, MMAE-based model adaptation, SVM-based classification, and PPO-based control optimization. The UKF and MMAE components involve recursive estimation with complexity proportional to the state dimension and number of candidate models, while the SVM classifier operates efficiently during inference with moderate computational overhead. The PPO agent introduces additional training complexity; however, its inference phase remains suitable for real-time deployment. Overall, the modular structure allows parallel processing and efficient implementation, making the framework scalable and feasible for large-scale industrial applications with appropriate computational resources.

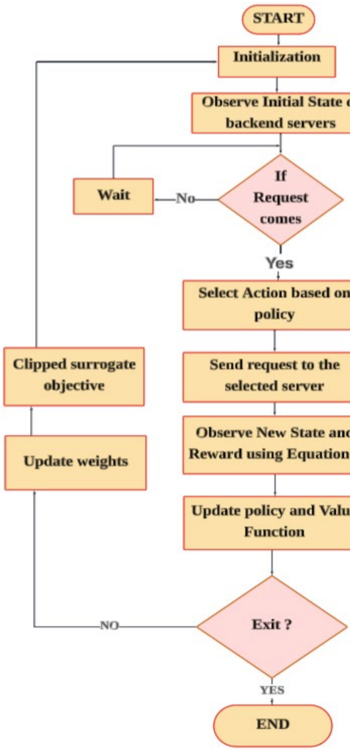

Figure 3. Proximal Policy Optimization Flowchart

4. Result

This study has demonstrated that the proposed framework has very effectively combined anomaly detection and adaptive control in intelligent manufacturing systems. The experiments were conducted with Python 3.11.9 on a Windows 10 (64-bit) operating system. For data processing, modeling, and visualization, the libraries NumPy (1.26.4), Pandas (2.3.1), Scikit-learn (1.7.2), and Matplotlib (3.10.3) were used along with PyCharm and Jupyter Notebook as development tools. The hardware platform included an Intel® CPU with 12 cores and 8.32 GB RAM, and all computations were carried out on the CPU with no GPU acceleration. The SVM classifier reached a perfect ROC–AUC of 1.000, vividly illustrating the difference between normal and abnormal states, while the residual-based analysis indicates that the threshold crossings during faulty conditions and the stable behavior otherwise are indeed separate. The merging of the sensors by the UKF and MMAE not only ensures accuracy in estimation but also provides more robustness during the dynamic conditions. The award of the PPO-based DRL controller goes up consistently and the control effectiveness remains stable, which means that policy learning and adaptive decision-making have taken place. The PPO agent underwent training for 3000 episodes with a batch size set to 64, a learning rate to 3×10^{-4}, a discount factor to 0.99, and a clipping parameter to 0.2. The SVM classifier used an RBF kernel with a regularization parameter of C = 1.0, and overfitting was controlled through regularization without the use of dropout, as dropout is not applicable to SVM-based models.

4.1 Performance Metrics

In order to carry out a quantitative assessment of the framework's usefulness, the standard performance metrics for anomaly detection, state estimation, and control performance are considered. The metrics for anomaly detection are Accuracy, Precision, Recall, F1-score, and Receiver Operating Characteristic–Area Under Curve (ROC–AUC). The definitions of these metrics are as follows:

Accuracy measures the overall correctness of the model shown in Eqn. (20).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (20)$$

Precision measures the proportion of correctly predicted positive samples shown in Eqn. (21).

$$\text{Precision} = \frac{TP}{TP+FP} \quad (21)$$

Recall (Sensitivity) measures the ability of the model to identify positive cases shown in Eqn. (22).

$$\text{Recall} = \frac{TP}{TP+FN} \quad (22)$$

F1-score is the harmonic mean of Precision and Recall shown in Eqn. (23).

$$\text{F1-score} = \frac{2\times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (23)$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

The ROC–AUC metric measures the trade-off between the true positive rate and the false positive rate, and it is utilized to evaluate the overall discrimination power of the anomaly detection model.

In the control evaluation, the PPO agent's cumulative reward per time step is applied, which is defined as shown in Eqn. (24).

$$\mathrm{R} = \sum_{t=0}^{T} \gamma^{t} r_{t} \quad (24)$$

where r_t is the reward at time step t, T is the episode length, and γ is the discount factor. Higher reward values indicate better adaptive control performance and system stability.

4.2 ROC Curve of SVM Classifier Showing Perfect Anomaly Separation

The ROC curve illustrates the classification capability of the SVM model, which achieved a perfect AUC score of 1.000, indicating total separation between normal and anomalous states shown in Figure 4. The curve comes extremely close to the top-left corner, which indicates very high sensitivity and specificity. The model has surpassed the random baseline significantly with regard to predictive power. This result not only proves the feature vector to be effective but also gives credence to the anomaly classification step in the proposed control framework.

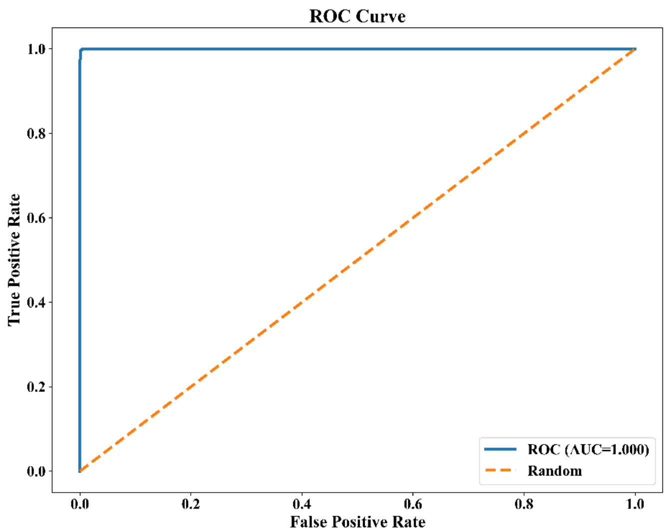

Figure 4. SVM-Based Anomaly Detection Accuracy with AUC Benchmarking

4.3 Residual-Based Anomaly Detection Analysis Using Time-Series, Scatter, and Distribution Metrics

The performance of the anomaly detection method based on residuals in the proposed framework is demonstrated through various visual analyses as shown in the figures. In the Figure 5, the time-series visualization of anomaly severity using the residual magnitude with the threshold benchmarking is shown, where the residual values remain within the acceptable bounds during normal operation but show sharp spikes that exceed the predefined threshold during the abnormal conditions. The threshold crossings in this way designate periods of major deviation in the system quite clearly. Figure 6 gives additional proof to the detection mechanism by means of a scatter plot of residuals, which normal operating points concentrate below the threshold while anomalous cases are distinctly separated above it, hence showing effective real-time anomaly discrimination. In Figure 7, the temporal behavior of the residuals is again investigated, wherein the residual magnitude over time is depicted showing that there are frequent low-level fluctuations that are interspersed with prominent spikes that correspond to the fault occurrences. Also, the frequency distribution of the residual magnitudes illustrated in Figure 8 is revealing a right-skewed pattern, where most of the residuals are concentrated on the lower magnitudes indicating stable system behavior, whereas the high-magnitude residuals occur rarely and are correlated to the anomalous events. All these figures together testify to the strength, reliability, and intelligibility of the residual-based anomaly detection method for continuous monitoring in intelligent manufacturing systems.

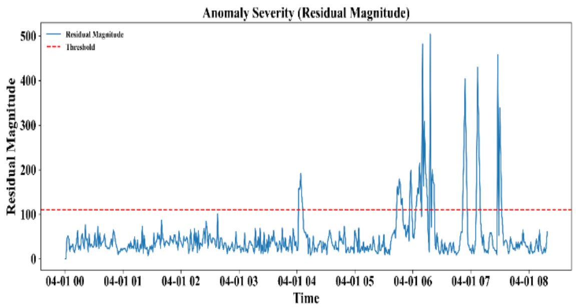

Figure 5. Time-Series Visualization of Anomaly Severity via Residual Magnitude and Threshold Benchmarking

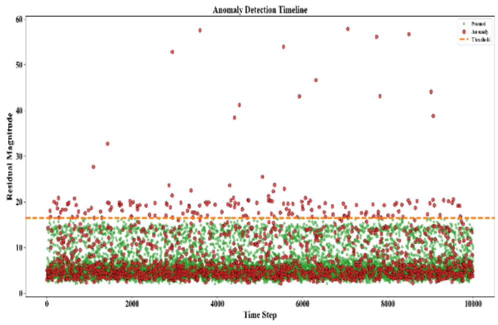

Figure 6. Scatter Plot of Residual Magnitudes for Real-Time Anomaly Identification

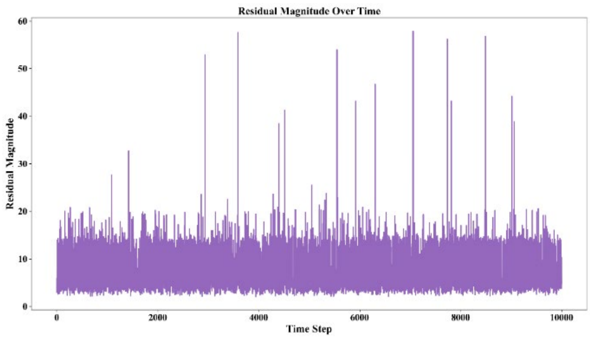

Figure 7. Time-Series Analysis of Residual Magnitude in Smart Manufacturing System

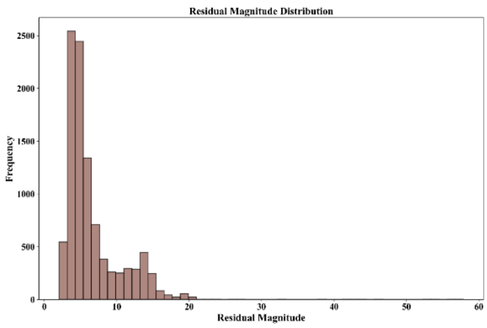

Figure 8. Frequency Distribution of Residuals for Model Performance Assessment

4.4 Sensor Signal Behavior and Data Quality

This diagram shows the time series data of four sensors across 10,000-time steps, highlighting the different behaviors of the signals shown in Figure 9. SENSOR_00 has fast changes but within a small amplitude, whereas SENSOR_01 is stable but there is a sharp drop near time step 6000. SENSOR_10 is quite variable and sometimes it goes down to zero for a while. SENSOR_20 is always high, has dense, and very brief drops. The stated signals can be utilized for recognizing the sensors that work the best, to monitor the quality of the signals as well as to detect the possible abnormalities in the system.

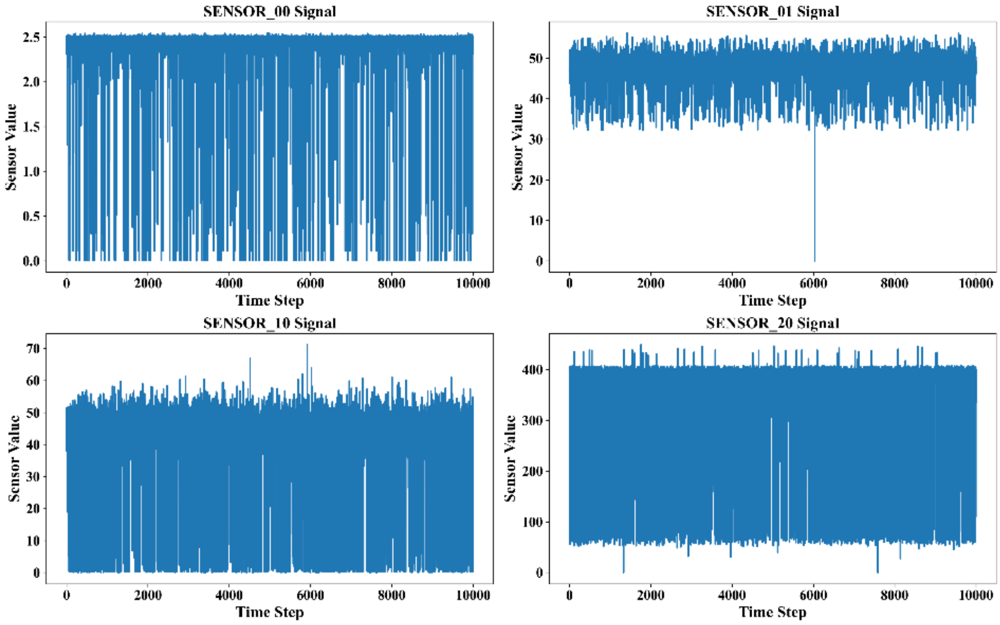

Figure 9. Sensor Signal Dynamics Across Four Channels in Smart Manufacturing Environment

4.5 Evaluation of UKF–MMAE-Based State Estimation Accuracy Across Four State Dimensions State Estimation Performance (UKF vs MMAE)

This Figure 10 illustrates the state estimation performance comparison of the Original UKF and MMAE Refined methods through 10,000 time steps and four state dimensions. The subplots depict a very close alignment of the solid blue (UKF) and the dashed red (MMAE) lines, which means that the refined MMAE method is capable of following the original UKF estimates closely. The similarity through all dimensions points to the fact that the MMAE refinement is a method that retains estimation accuracy while at the same time routing the possibly improved robustness or adaptability in dynamic environments.

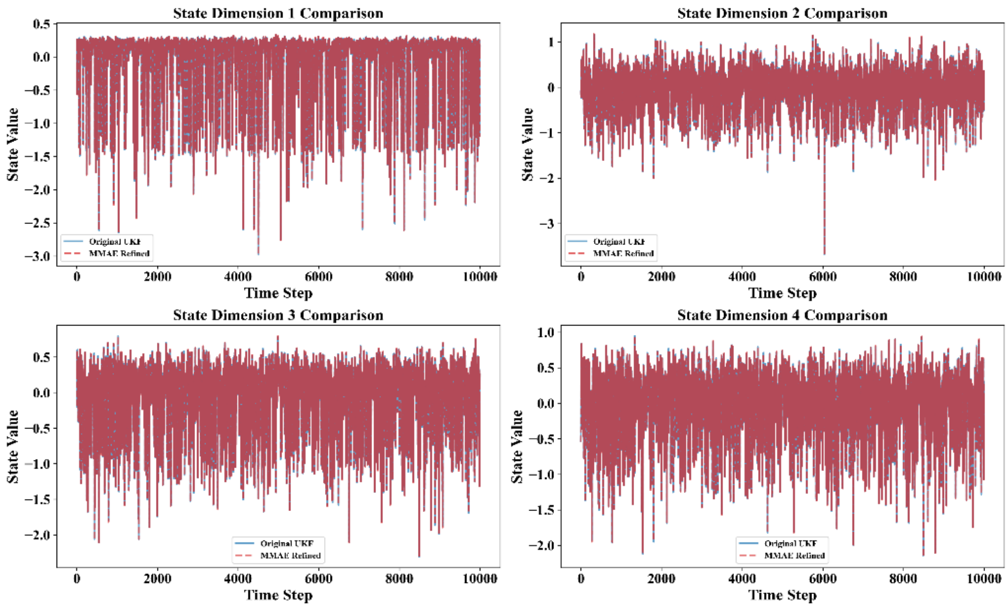

Figure 10. State Estimation Fidelity of UKF–MMAE in Smart Manufacturing Control

4.6 DRL Training Performance

The graph showcases the PPO algorithm’s reward evolution throughout 3000 training episodes. The dense purple line is a representation of unstable reward values, shown in Figure 11, the dynamic learning behavior of the agent. The red dashed line represents the average reward at 1.793 which is an indication of policy improvement happening consistently. Such a result demonstrates that the PPO agent was capable of continuously optimizing its control actions, which led to increased efficiency in production and reduced anomaly detection. The upward reward trend is proof of the successful implementation of the DRL-based adaptive control system in manufacturing processes.

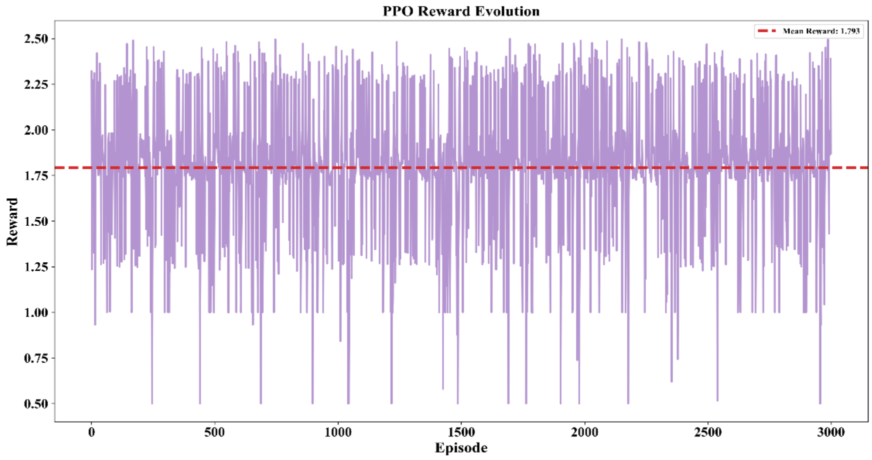

Figure 11. Residual Signal Behavior for Real-Time Fault Detection

4.7 DRL Control Effectiveness

The control effectiveness of the system based on DRL is depicted by the graph for a time range of 3000 steps shown in Figure 12. The green line continues to indicate different performance levels, which are the outcomes of the responsive nature of the system to the changes in its surroundings. A red dashed line has been drawn to represent the average effectiveness at 0.775, which shows that the system had very good control behavior most of the time. The range for control effectiveness is set to [0, 1], with numbers exceeding 0.7 becoming suitable for stable industrial control. The DRL-based control algorithm's average effectiveness of 0.775 indicates that the performance is very good even under changing operation conditions. The adaptation of the system, in spite of the differences, remains quite uniform. The result confirms that the policy gradient agent possesses the ability to carry out control actions that are stable and at the same time efficient in intelligent manufacturing settings.

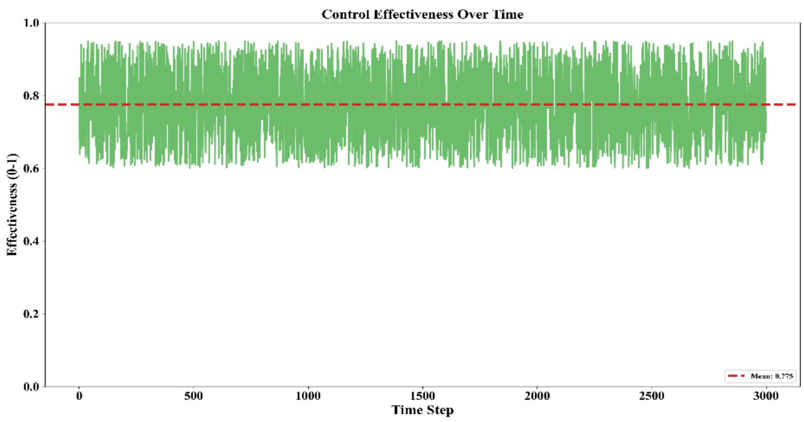

Figure 12. Temporal Stability of DRL-Based Control Effectiveness Over 3000 Steps

4.8 DRL Action Selection Behavior

The illustration shows how the likelihoods of selecting the action changed over 3000 time steps for the different control actions which are Speed Adjustment, Temperature Control and Feed Rate shown in Figure 13. The green line which represents the Feed Rate shows the highest probability all the time indicating that it was most often the preferred choice of the DRL agent. Next in line is Speed Adjust, while Temp Control is still the least favored option. The variations in these percentages can be interpreted as the agent's use of adaptive strategies which were determined by the present business conditions and the rewards it earned; hence, the Doe PPO agent's skill of operating through changing manufacturing setups was verified.

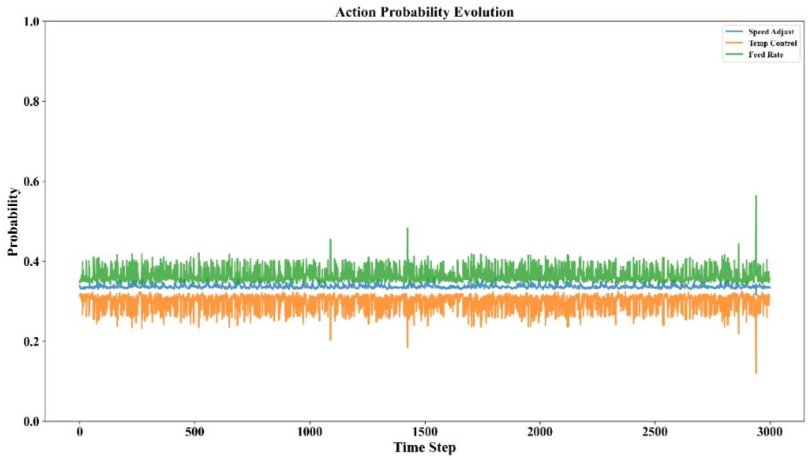

Figure 13. Probability Shifts in Adaptive Control Decisions During Smart Manufacturing Operations

4.9 Integrated DRL State Dynamics

The Figure 14 illustrates the gradual changes in a sequential decision system over 3000 steps. The continuous state subplot maps out the internal fluctuations, while the residual magnitude spots the times of maximum error or uncertainty. The classification labels do not change very much, which means that the detection is stable and the past actions correspond to a fixed policy. The three plots give us a brief yet comprehensive picture of the system's adaptation, discrepancy evaluation, and response through classification and action over time.

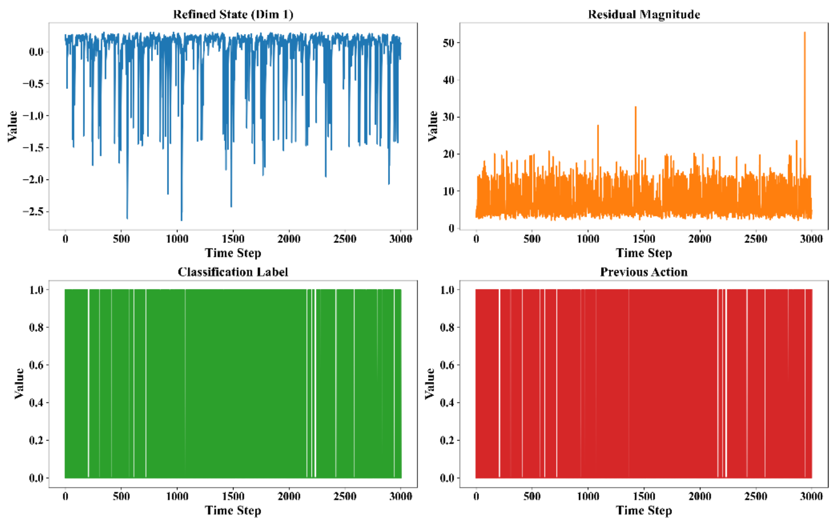

Figure 14. DRL State Construction Components Over Time in Smart Manufacturing Control

5. Discussion

The experimental results demonstrate the effectiveness of the proposed intelligent manufacturing framework that integrates multi-sensor fusion, residual-based anomaly detection, and deep reinforcement learning-based adaptive control. The combination of Unscented Kalman Filter (UKF) and Multi-Model Adaptive Estimation (MMAE) provides accurate and robust state estimation under nonlinear and uncertain operating conditions. The close alignment between UKF and MMAE-refined estimates across all state dimensions indicates that the refinement process enhances system reliability without compromising estimation accuracy.

The residual-based anomaly detection mechanism effectively captures deviations between observed and estimated sensor values, enabling timely identification of abnormal operating conditions. The Support Vector Machine (SVM) classifier further improves the reliability of anomaly classification by clearly separating normal and faulty states, as evidenced by the high ROC–AUC performance.

The integration of anomaly information and fused state estimates into the Proximal Policy Optimization (PPO)-based deep reinforcement learning agent enables adaptive and real-time control decisions. The observed improvement in cumulative rewards and stable control effectiveness confirms that the DRL agent successfully learns optimal policies under dynamic manufacturing conditions.

Despite the strong performance, certain limitations remain. The evaluation is conducted on a publicly available dataset, and further validation on large-scale real industrial systems is necessary. Additionally, the framework assumes reliable sensor availability, and future work may explore fault-tolerant learning and multi-agent reinforcement learning for enhanced scalability. Overall, the proposed UKF–MMAE-based pipeline demonstrates strong potential for real-time anomaly-aware adaptive control in intelligent manufacturing environments.

6. Conclusion

This manuscript introduced a holistic deep reinforcement learning-based adaptive control system for smart manufacturing that merges multi-sensor fusion, residual-based anomaly detection, and robust state estimation in a closed-loop framework. The fusion of UKF-based sensor data with the MMAE process works through reliable state estimation even with nonlinear dynamics and uncertain sensors. If the residuals are analyzed by means of statistical thresholding, then one can detect the anomalies accurately, and SVM classification can help in distinguishing the faults. All these components brought together in a PPO-based deep reinforcement learning controller, the system can modify control actions promptly to the changes in detected anomalies and operating conditions. The experiments done proved the proposed method's effectiveness showing high anomaly detection accuracy, learning stability, and control performance consistency over a long operation period. In short, the offered adaptive control framework boosts the reliability, robustness, and flexibility of intelligent manufacturing and is therefore an attractive option for the future smart factories. Next steps will include trying it out in real industrial situations, making it applicable to large-scale production, and integrating safety-aware and multi-agent reinforcement learning approaches.

Declarations

Conflict of Interest

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

Author Contributions

Shanshan Kong contributed to the conceptualization, methodology design, data preprocessing, model implementation, experimental analysis, and manuscript drafting.
Peng Zhou contributed to supervision, validation, results interpretation, critical revision of the manuscript, and final approval of the submitted version.
All authors have read and agreed to the published version of the manuscript.

References

[1] R. Sahba, R. Radfar, A. Rajabzadeh Ghatari, and A. Pour Ebrahimi, “Development of Industry 4.0 predictive maintenance architecture for broadcasting chain,” *Adv. Eng. Inform.*, vol. 49, p. 101324, Aug. 2021, doi: 10.1016/j.aei.2021.101324.

[2] V. Varriale, A. Cammarano, F. Michelino, and M. Caputo, “Critical analysis of the impact of artificial intelligence integration with cutting-edge technologies for production systems,” *J. Intell. Manuf.*, vol. 36, no. 1, pp. 61–93, Jan. 2025, doi: 10.1007/s10845-023-02244-8.

[3] Q. Lawal, "Leveraging Artificial Intelligence to Enhance Process Control and Improve Efficiency in Manufacturing Industries," *Int. J. Comput. Appl. Technol. Res.*, vol. 14, no. 2, Jan. 2025, doi: 10.7753/IJCATR1402.1002.

[4] D. Frempong *et al.*, "Real-Time Analytics Dashboards for Decision-Making Using Tableau in Public Sector and Business Intelligence Applications," *J. Front. Multidiscip. Res.*, vol. 3, pp. 65–80, Jan. 2022, doi: 10.54660/.IJFMR.2022.3.2.65-80.

[5] M. A. Rawajbeh *et al.*, "Trustworthy Adaptive AI for Real-Time Intrusion Detection in Industrial IoT Security," *IoT*, vol. 6, no. 3, Sep. 2025, doi: 10.3390/iot6030053.

[6] A. Siriweera and K. Naruse, "QoS-Aware Federated Crosschain-Based Model-Driven Reference Architecture for IIoT Sensor Networks in Distributed Manufacturing," *IEEE Sens. J.*, vol. 23, no. 23, pp. 29630–29644, Dec. 2023, doi: 10.1109/JSEN.2023.3325342.

[7] G. Mattera, A. Caggiano, and L. Nele, "Optimal data-driven control of manufacturing processes using reinforcement learning: an application to wire arc additive manufacturing," *J. Intell. Manuf.*, vol. 36, no. 2, pp. 1291–1310, Feb. 2025, doi: 10.1007/s10845-023-02307-w.

[8] I. Podder, T. Fischl, U. Bub, I. Podder, T. Fischl, and U. Bub, "Artificial Intelligence Applications for MEMS-Based Sensors and Manufacturing Process Optimization," *Telecom*, vol. 4, no. 1, pp. 165–197, Mar. 2023, doi: 10.3390/telecom4010011.

[9] F. J. Maseda *et al.*, "Sensors Data Analysis in Supervisory Control and Data Acquisition (SCADA) Systems to Foresee Failures with an Undetermined Origin," *Sensors*, vol. 21, no. 8, Apr. 2021, doi: 10.3390/s21082762.

[10] W. Strielkowski *et al.*, "Prospects and Challenges of the Machine Learning and Data-Driven Methods for the Predictive Analysis of Power Systems: A Review," *Energies*, vol. 16, no. 10, May 2023, doi: 10.3390/en16104025.

[11] F. Ale, I. Daniyan, O. Aderoba, and A. A. Adediran, "An Artificial Intelligence-based Integrated Framework for Lean and Smart Manufacturing: A Case Study of the Rail Industry," in *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG)*, Omu-Aran, Nigeria: IEEE, Apr. 2024, pp. 1–10. doi: 10.1109/SEB4SDG60871.2024.10629752.

[12] A. Corallo, M. Lazoi, M. Lezzi, and P. Pontrandolfo, "Cybersecurity Challenges for Manufacturing Systems 4.0: Assessment of the Business Impact Level," *IEEE Trans. Eng. Manag.*, vol. 70, no. 11, pp. 3745–3765, Nov. 2023, doi: 10.1109/TEM.2021.3084687.

[13] L. Rojas, Á. Peña, J. Garcia, L. Rojas, Á. Peña, and J. Garcia, "AI-Driven Predictive Maintenance in Mining: A Systematic Literature Review on Fault Detection, Digital Twins, and Intelligent Asset Management," *Appl. Sci.*, vol. 15, no. 6, Mar. 2025, doi: 10.3390/app15063337.

[14] B. Xia, K. Wang, A. Xu, P. Zeng, N. Yang, and B. Li, "Intelligent Fault Diagnosis for Bearings of Industrial Robot Joints Under Varying Working Conditions Based on Deep Adversarial Domain Adaptation," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–13, 2022, doi: 10.1109/TIM.2022.3158996.

[15] J. Hu, C. Fan, M. Ozay, Q. Gao, Y. Guo, and T. L. Lam, "Robust Depth Estimation Under Sensor Degradations: A Multi-Sensor Fusion Perspective," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 47, no. 10, pp. 8691–8707, Oct. 2025, doi: 10.1109/TPAMI.2025.3581311.

[16] M. Shahin, A. Hosseinzadeh, and F. F. Chen, "Generative artificial intelligence in manufacturing: applications, case studies, and future directions for next-generation intelligent production systems," *Int. J. Adv. Manuf. Technol.*, vol. 141, no. 3, pp. 1159–1265, Nov. 2025, doi: 10.1007/s00170-025-16667-5.

[17] Z. Lian, Z. Zhou, C. Hu, Z. Ming, J. Wang, and Y. Zhao, "A Belief Rule-Based Performance Evaluation Model for Complex Systems Considering Sensors Disturbance," *IEEE Trans. Reliab.*, vol. 73, no. 2, pp. 1245–1257, Jun. 2024, doi: 10.1109/TR.2023.3311436.

[18] D. A. Bhanage, A. V. Pawar, and K. Kotecha, "IT Infrastructure Anomaly Detection and Failure Handling: A Systematic Literature Review Focusing on Datasets, Log Preprocessing, Machine & Deep Learning Approaches and Automated Tool," *IEEE Access*, vol. 9, pp. 156392–156421, 2021, doi: 10.1109/ACCESS.2021.3128283.

[19] J. Desikan *et al.*, "Hybrid Machine Learning-Based Fault-Tolerant Sensor Data Fusion and Anomaly Detection for Fire Risk Mitigation in IIoT Environment," *Sensors*, vol. 25, no. 7, Mar. 2025, doi: 10.3390/s25072146.

[20] X. Xiao, J. Dufek, and R. R. Murphy, "Robot Risk-Awareness by Formal Risk Reasoning and Planning," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 2856–2863, Apr. 2020, doi: 10.1109/LRA.2020.2974434.

[21] M. Orabi, K. P. Tran, P. Egger, and S. Thomassey, "Anomaly detection in smart manufacturing: An Adaptive Adversarial Transformer-based model," *J. Manuf. Syst.*, vol. 77, pp. 591–611, Dec. 2024, doi: 10.1016/j.jmsy.2024.09.021.

[22] C.-C. Wong, H.-M. Feng, and K.-L. Kuo, "Multi-Sensor Fusion Simultaneous Localization Mapping Based on Deep Reinforcement Learning and Multi-Model Adaptive Estimation," *Sensors*, vol. 24, no. 1, p. 48, Dec. 2023, doi: 10.3390/s24010048.

[23] T. Zhou, M. Chen, and J. Zou, “Reinforcement Learning Based Data Fusion Method for Multi-Sensors,” *IEEECAA J. Autom. Sin.*, vol. 7, no. 6, pp. 1489–1497, 2020, doi: 10.1109/JAS.2020.1003180.
[24] Zhe Sun, Minru Kong, and Guiqi Zhu, “Multi-source Sensor Data Fusion and Deep Reinforcement Learning Anomaly Detection Method for Power UAV Inspection,” *Int. J. Hous. Sci. Its Appl.*, vol. 46, no. 3, 2025.
[25] Y. Yu and M. Liu, “Deep Reinforcement Learning-Based Multisensor Control for Labeled Multi-Bernoulli Filtering,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 61, no. 5, pp. 13548–13564, Oct. 2025, doi: 10.1109/TAES.2025.3579476.
[26] Y. M. Alginahi, O. Sabri, and W. Said, “Reinforcement Learning for Industrial Automation: A Comprehensive Review of Adaptive Control and Decision-Making in Smart Factories,” *Machines*, vol. 13, no. 12, p. 1140, Dec. 2025, doi: 10.3390/machines13121140.
[27] R. Anwar, J.-W. Kwon, and W.-T. Kim, “A Deep Reinforcement Learning-Based Concurrency Control of Federated Digital Twin for Software-Defined Manufacturing Systems,” *Appl. Sci.*, vol. 15, no. 15, p. 8245, Jul. 2025, doi: 10.3390/app15158245.
[28] J. Huang *et al.*, “Deep Reinforcement Learning-Based Dynamic Reconfiguration Planning for Digital Twin-Driven Smart Manufacturing Systems With Reconfigurable Machine Tools,” *IEEE Trans. Ind. Inform.*, vol. 20, no. 11, pp. 13135–13146, Nov. 2024, doi: 10.1109/TII.2024.3431095.