Czech Technical University in Prague
Faculty of Mechanical Engineering

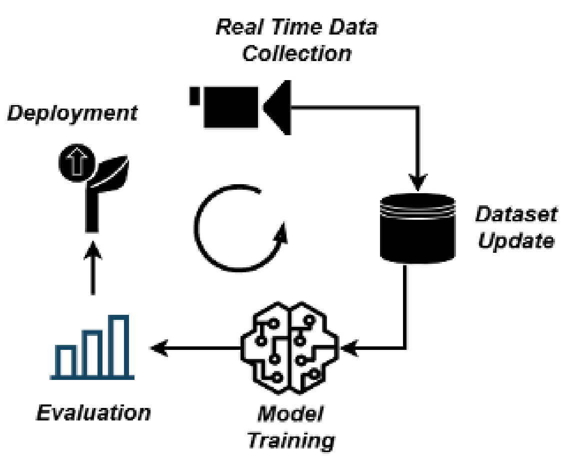
Project SGS24/126/OHK2/3T/12

Report of 2025

1. Abstract and Core Innovation

This research presents the design and empirical validation of an advanced, embedded predictive quality-control (QC) system for Fused Deposition Modeling (FDM) processes. The system's central innovation is a critical shift from reactive post-hoc defect detection to proactive, constraint-aware defect prevention. This is achieved via a tightly coupled, closed-loop control system that integrates four distinct, critical layers: Multi-Sensor Process Monitoring, a compact Defect Probability Estimation (DPE) model, Risk-Gated Vision Anomaly Verification (VAV) via a Convolutional Neural Network (CNN), and a Tube-Based Robust Model Predictive Control (MPC) engine. The entire system is deployed on low-cost, embedded hardware (Raspberry Pi 5) and is designed to operate within a stringent 200 ms real-time control cycle (5 Hz). Crucially, the system requires no modifications to the base Marlin firmware, communicating exclusively via safe, non-blocking G-code. Experimental validation across 1000 print jobs demonstrated a 60% reduction in the defective print rate, conclusively validating the feasibility and economic viability of this advanced, deterministic control framework in desktop additive manufacturing.

This system relies on an Adaptive learning Scheme [1], which allows for the continuous refinement of the predictive models over time.



[1] Adaptive learning Scheme

2. Project Objectives: A Focus on Deterministic Control

The core objectives of this project focused on developing a technically advanced, yet practical and robust, control system:

- Predict Defects Before They Manifest: Exploit multi-sensor process data to estimate defect risk earlier than is possible with traditional vision-based inspection.
- Minimize Computational Load via Risk-Gated Vision: Activate computationally expensive CNN inference (for VAV) only when the predicted risk from the DPE model justifies the additional computational load.
- Apply Robust, Constraint-Aware Control: Use tube-based MPC to ensure safety, stability, and feasibility of control actions even under system disturbances.
- Achieve Fully Embedded Real-Time Operation: Maintain deterministic timing within a strict 200 ms control cycle on the low-cost, embedded hardware platform.
- Provide Full System Transparency: Develop a dedicated Python dashboard for live monitoring, logging of critical events, and detailed post-process analysis.
- Ensure Replicability and Practical Deployment: Rely exclusively on off-the-shelf sensors, open-source software, and standard printer interfaces for ease of adoption.

3. Background and Rationale

3.1 Industrial and Research Context

FDM 3D printing is widely adopted but remains highly sensitive to process drift, material variability, and environmental conditions. These instabilities frequently lead to common defects such as under-extrusion, stringing, blobs, warping, and cracking, which compromise part quality and result in wasted time and material. Current QC approaches fall short: manual inspection is subjective and post-hoc; rule-based thresholds lack adaptability; and continuous vision-based detection is reactive and computationally expensive. This system addresses the need for early, proactive, and embedded defect prevention.

3.2 Technical Motivation

This work was motivated by three key technical observations:

1. Process sensors reveal anomalies earlier than vision: Signals like temperature, force, and filament feed irregularities often provide indications that precede visible defects.
2. CNNs are powerful but costly: Running continuous, high-fidelity vision inference wastes compute resources when the print is operating healthily.

3. MPC naturally enforces safety and anticipation: Model Predictive Control (MPC) is highly suitable for constrained systems like FDM printers, where anticipating future states and avoiding unsafe actions is critical.

These insights were combined to form a single, unified predictive and corrective control framework.

4. System Overview

The system is built upon four tightly coupled layers that form the closed-loop control structure:

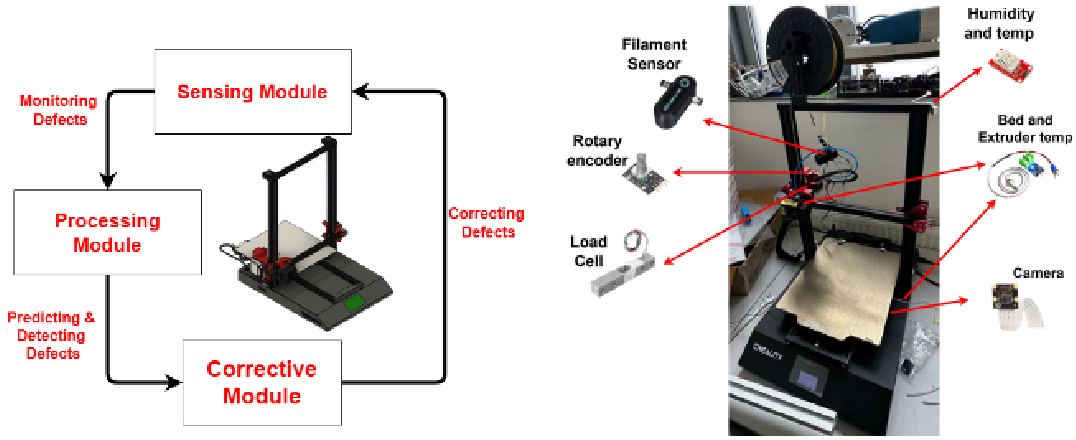
1. Multi-Sensor Monitoring
2. Defect Probability Estimation (DPE)
3. Risk-Gated Vision Anomaly Verification (VAV)
4. Robust Model Predictive Control (MPC)

The entire architecture is spanned by a Python dashboard that provides live observability across all layers.

5. System Architecture and Methodology

5.1 Hardware and Sensor Instrumentation

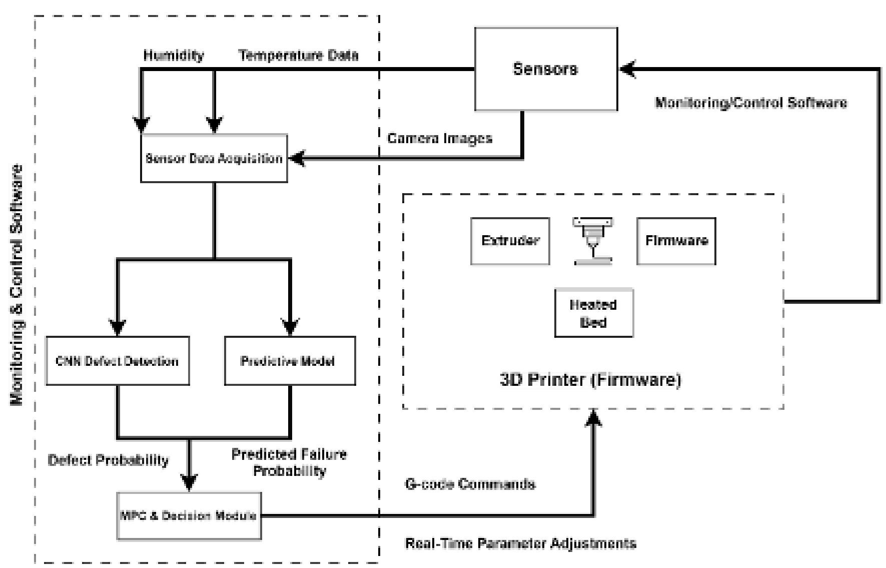
A standard desktop FDM printer (Prusa i3 class) was instrumented with a suite of low-cost sensors: nozzle and bed temperature, extrusion force, filament motion, ambient temperature and relative humidity, and an RGB camera for visual inspection. All sensor signals are precisely time-aligned and sampled at 5 Hz to match the required control cycle frequency. The implementation demonstrates a robust, yet low-cost, control solution, High-End System [2] is an overview with three modules that work together to achieve the purpose of this project.



[2] High-End System

5.2 Feature Engineering and Defect Probability Estimation

Five primary process variables are monitored. To create a feature set that captures short-term trends without imposing heavy memory usage, a 5-second exponentially weighted moving average (EWMA) is calculated and appended for each of the five channels, forming a 10-dimensional feature vector for each control cycle. This vector is mapped to a defect probability using a logistic regression model, which was selected for its high interpretability, deterministic inference time, and numerical stability on embedded hardware. System Overview [3] illustrates how these components interact.



[3] System Overview

5.3 Risk-Gated Vision Anomaly Verification

The vision module is conditionally activated. When the predicted defect probability from the DPE model exceeds a calibrated threshold, a YOLOv4-tiny CNN is executed on the latest camera frame. The vision output is then used to refine the risk estimate, utilizing a conservative fusion strategy to ensure system safety and reliability. This gating strategy is vital for reducing unnecessary CNN execution, thereby conserving precious computational resources while preserving the integrity of the detection reliability.

5.4 Robust Model Predictive Control

When the fused defect risk (DPE + VAV) remains high, the system engages the **tube-based robust MPC**. This controller is built upon a linear plant model identified from experimental data. It handles uncertainty by explicitly bounding disturbances and tightening control constraints using formally derived invariant sets. The optimization problem is solved efficiently using the OSQP solver with warm-starting. The controller's output is designed to be **small, incremental, and safe parameter adjustments**, explicitly preventing aggressive or potentially unsafe interventions that could destabilize the printing process. **Tube-based MPC concept with nominal trajectory and disturbance bounds [4]** explains the principle of the robust control method.

Quantity	Value	Note
Nozzle model R^2	0.96 ± 0.01	step/PRBS fits
Extrusion rate R^2	0.91 ± 0.02	$\dot{m}$ dynamics
$\|A\|_2$	0.87 ± 0.03	stable nominal
$\|B\|_2$	0.41 ± 0.05	input gain magnitude
Disturbance bound $\bar{w}$	0.8	95% quantile (z-scored)

[4] Tube-based MPC concept with nominal trajectory and disturbance bounds.

5.5 Software Workflow and Real-Time Execution

The entire software pipeline is fully non-blocking and executes the following sequence within the 200 ms loop: 1. Sensor acquisition; 2. Feature update and DPE inference; 3. Conditional CNN inference; 4. MPC optimization (if gated); 5. Safe G-code transmission; 6. Logging and dashboard update. Only whitelisted, non-blocking G-codes are issued to the printer. This full process is shown in the Real-Time Workflow [5].

Algorithm 1: Monitoring & Control Loop (200 ms cycle; fused-risk gating, non-blocking I/O)

Input: sensor vector $\mathbf{x}_k$, RGB frame I_k

Output: optional G-code correction vector $\Delta\mathbf{u}$

$P_d \leftarrow \sigma(\boldsymbol{\beta}^\top \boldsymbol{\phi} + b)$;

if $P_d \leq \tau$ **then**

  $\hat{P}_d \leftarrow P_d$;

else

  $\hat{y} \leftarrow \mathrm{VAV}(I_k)$; $\hat{P}_d \leftarrow \max(P_d, \hat{y})$;

if $\hat{P}_d > \tau$ **then**

  if `solve_MPC`($\hat{P}_d$, *timeout=20 ms*) **then**

    $\Delta\mathbf{u} \leftarrow$ first_action();

  else

    $\Delta\mathbf{u} \leftarrow$ fallback_K($\mathbf{x}_k - \hat{\mathbf{x}}_k$);

    (actuation suppressed this cycle)

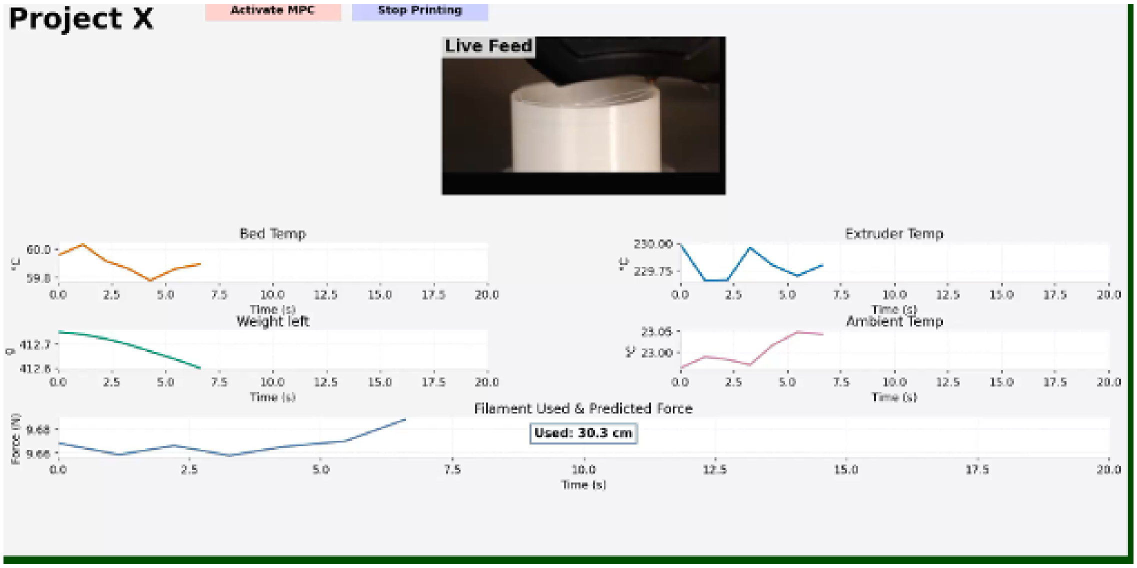
  `enqueue_gcode_nonblocking`($\Delta\mathbf{u}$, drop_oldest=true);

`log`($\mathbf{x}_k, P_d, \hat{P}_d, \Delta\mathbf{u}$);

[5] Real-Time Workflow

6. Python Monitoring Dashboard

A custom Python dashboard was developed to provide comprehensive, real-time insight into the system's behavior. It visualizes: live sensor signals, defect probability and vision confidence, MPC outputs and safety limits, system latency and CPU usage, and event logs with defect markers. The dashboard functions as both an operator interface and a research tool for dataset generation, debugging, and model refinement. Python dashboard showing live monitoring and control signals [6] is an example of the information provided to the user.



[6] Python dashboard showing live monitoring and control signals

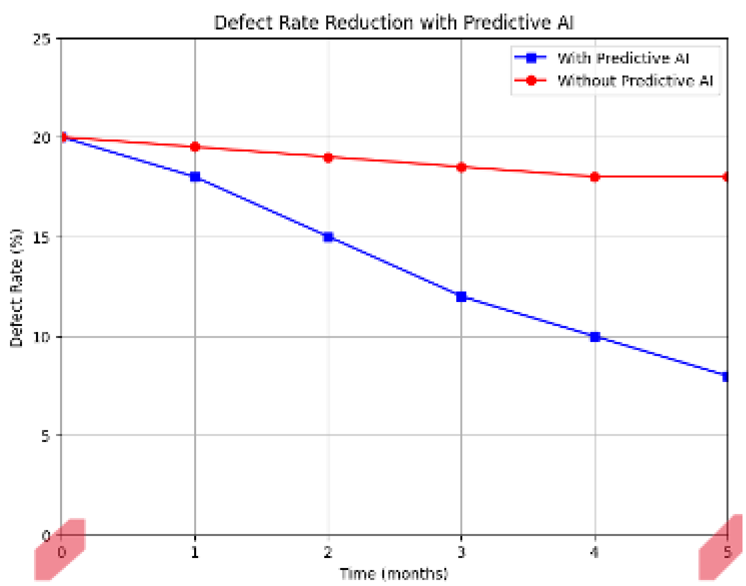
7. Experimental Setup

The physical testing involved a Desktop FDM printer (Prusa i3 class) using PLA material in a controlled temperature and humidity environment. A total of 1000 print jobs were evaluated at the target 5 Hz control frequency. Ground truth labels were established via a rigorous combination of human review and high-confidence CNN annotation for objective performance measurement.

8. Results and Evaluation

The system's impact was quantifiable and significant:

- The baseline defect rate was reduced from 20 % to 8 % (a 60% relative reduction).
- The percentage of flawless prints increased substantially, from 65 % to 88 %.
- The combined predictive accuracy of the DPE/VAV system reached ≈ 91 %.
- The mean control loop latency was measured at 132 ms, well below the 200 ms budget, validating the embedded design.
- Comparison of print outcomes with and without predictive control [7] graphically demonstrates this reduction in defect rate.



Metric	Without AI (%)	With AI (%)
Defect Rate	20.0	8.0
Reliability	65.0	88.0
False Positives	10.0	4.5
False Negatives	15.0	7.0

[7] Comparison of print outcomes with and without predictive control.

8.2 Ablation and Robustness Studies

Critical component studies revealed:

- Removing the vision (VAV) component caused the largest accuracy drop.
- Removing extrusion force data significantly reduced the system's predictive power.
- Tube-based MPC consistently outperformed a standard nominal MPC in terms of stability and reliability.

Model	Defect rate (%)		F1-score (%)	
	Mean	± CI	Mean	± CI
Threshold [6], [7]	19.4	0.8	61.0	2.0
YOLO-only	14.6	0.6	75.0	1.5
DPE-only (logistic)	12.9	0.5	78.0	1.2
SMC (no preview)	12.2	0.6	79.1	1.4
FOPID	11.6	0.5	80.3	1.3
Ours (no MPC)	10.1	0.4	83.0	1.0
Ours (nominal MPC)†	8.6	0.4	87.4	1.0
Ours (tube/robust MPC)	**8.0**	**0.3**	**88.0**	**0.9**

†MPC without tube tightening. Same folds and hyperparameters; tightening lowers defects from 8.6% to 8.0%.

[8] Comparison of print outcomes with and without predictive control.

The system demonstrated robustness across varying print speeds, humidity levels, and different PLA brands. The Comparison of print outcomes with and without predictive control [8] and Performance in different conditions [9] further illustrate these findings.

RH (%)	Defect rate (%)	Predictive accuracy (%)	Mean blobs
40	6.8	89.4	1.2
50	6.0	90.1	1.0
70	8.3	88.7	1.6

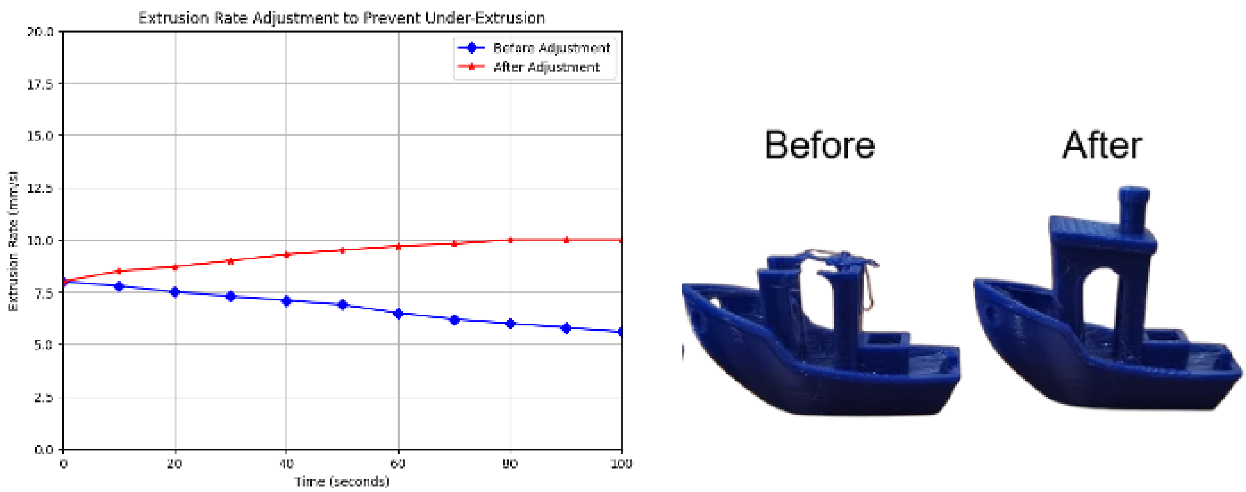
Speed ($\mathrm{mm\,s^{-1}}$)	Defect rate (%)	Predictive accuracy (%)	F1-score (%)
50	5.2	90.6	89.1
100	7.8	88.2	86.7
150	10.1	86.4	84.9

[9] Performance in different conditions (Humidity and Printing Speed)

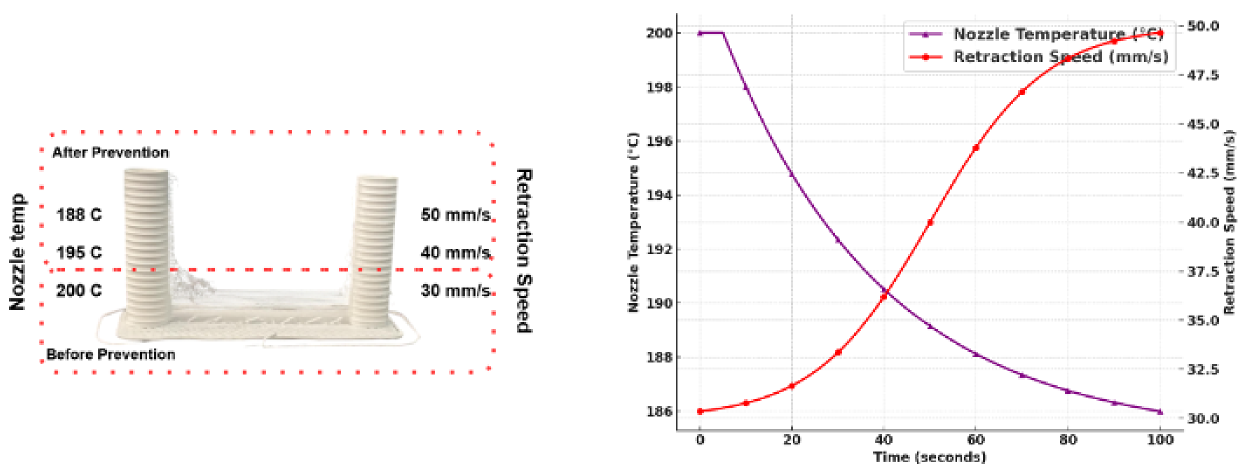
8.3 Case Studies

Specific control interventions were analyzed:

- Adaptive Extrusion Rate to Prevent Under-Extrusion [10] shows the system's ability to precisely modify material flow to counteract process drift that leads to under-extrusion.
- Retraction and Temperature Modifications to Minimize Stringing [11] illustrates coordinated, bounded parameter adjustments to prevent stringing defects.



[10] Adaptive Extrusion Rate to Prevent Under-Extrusion



[11] Retraction and Temperature Modifications to Minimize Stringing

9. Discussion

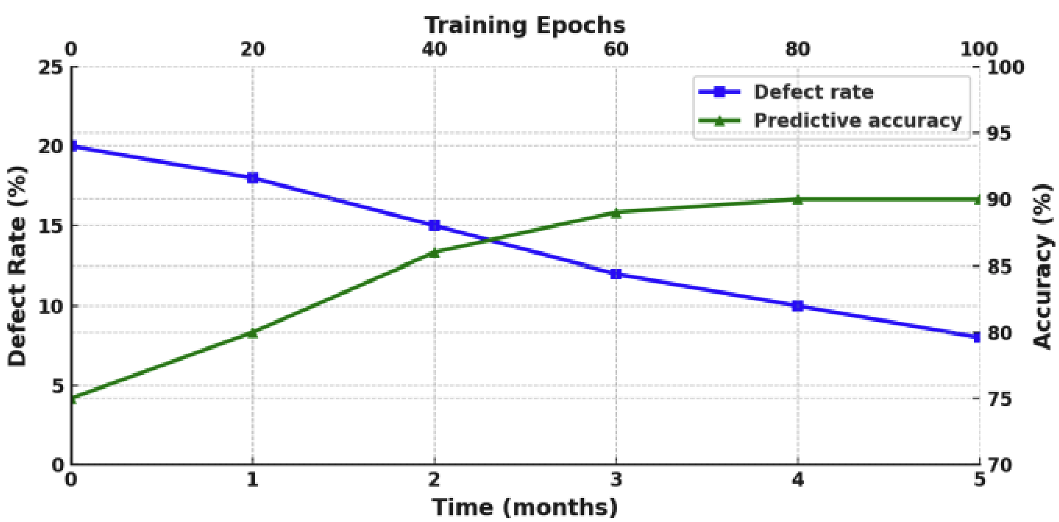
9.1 From Reactive Detection to Predictive Prevention

The key contribution of this work is the element of anticipation. By taking corrective action before defects fully manifest, the system avoids the need for abrupt, large-scale interventions, thereby preserving print continuity and structural integrity.

9.2 Embedded Intelligence at Low Cost

Successfully demonstrating robust MPC and CNN inference within a tight 200 ms loop on a commodity Raspberry Pi platform proves that advanced control and embedded intelligence are no longer limited to expensive, industrial-grade hardware, significantly lowering the barrier to entry for high-reliability additive manufacturing.

The Defect rate and predictive accuracy over time (five months) with monthly adaptive retraining [12] demonstrates the system's long-term stability and continuous improvement.



[12] Defect rate and predictive accuracy over time (five months) with monthly adaptive retraining

10. Conclusion and Future Work

10.1 Key Achievements

The project successfully delivered:

- Predictive, closed-loop defect prevention for FDM.
- Embedded implementation on low-cost hardware (Raspberry Pi 5).
- Robust control with formal safety guarantees (Tube-based MPC).
- Comprehensive monitoring via a custom Python dashboard.

10.2 Future Directions

Planned future work includes:

- Multi-camera and robotic camera integration for enhanced spatial monitoring.
- Adaptive MPC parameter learning to auto-tune the control system based on material and environmental changes.
- Extension to additional materials and printers beyond PLA and Prusa-class machines.
- Integration with predictive maintenance and digital twins for factory-level optimization.