Res Intell Manuf Assem, 2025, 4(1): 200-218
DOI: 10.25082/RIMA.2025.01.007

RESEARCH ARTICLE

Integrating Manufacturing Intelligence, Computer Vision, and Process Observation for Yield Improvement and Failure Prediction in Electronics Manufacturing

Vinit Vithalrai Shenvi[1] **Ashutosh Sharma**[1*]

[1] Department of Mechanical Engineering, M S Ramaiah Institute of Technology, Bengaluru, India

Correspondence to: Ashutosh Sharma, Department of Mechanical Engineering, M S Ramaiah Institute of Technology, Bengaluru, India;
Email: ashutosh.sharma0843@gmail.com

Received: March 30, 2025;
Accepted: June 13, 2025;
Published: June 19, 2025.

Citation: Shenvi VV, Sharma A. Integrating Manufacturing Intelligence, Computer Vision, and Process Observation for Yield Improvement and Failure Prediction in Electronics Manufacturing. *Res Intell Manuf Assem*, 2025, **4**(1): 200-218.
https://doi.org/10.25082/RIMA.2025.01.007

Copyright: © 2025 Vinit Vithalrai Shenvi et al. This is an open access article distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 International License, which permits all noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract: Electronics manufacturing processes are complex and prone to yield loss and latent failures due to subtle process deviations and quality escapes. This paper presents a holistic approach to improving first-pass yield and predicting failures by integrating a Manufacturing Intelligence for Reliability and Automated Insights (MIRAI) data platform with computer vision-based monitoring of Standard Operating Procedure (SOP) adherence. The proposed system combines self-serve data analytics workflows for yield and field failure analysis with real-time process observation using deep learning vision models. Manufacturing data from production tests, reliability screenings, and field returns are aggregated and analyzed to identify key signals correlating with yield drops and field fallouts. Simultaneously, a PROSPECT tool employs AI cameras at assembly stations to record operator actions and detect deviations from standard procedures. A machine learning failure prediction model is then trained on the enriched dataset (including vision-detected deviations) to proactively flag high-risk units in real time.

Keywords: computer vision, process observation, failure prediction, manufacturing analytics, yield improvement

1 Introduction

Modern electronics manufacturing is characterized by intricate assembly and test processes with thousands of interdependent steps. First-pass yield (FPY) is the percentage of units passing all tests without rework. Maintaining a high FPY is critical for cost efficiency and product quality [1]. Even minor deviations in how a process is executed can introduce defects that either cause immediate test failures or latent field failures after the product is shipped. Traditional quality control and yield analysis techniques often rely on reactive measures like identifying defects or yield drops only after they have occurred [1]. In many factories, engineers manually investigate yield excursions or perform periodic audits of operator compliance to SOPs (the standard work instructions for each task). These manual approaches are time-consuming, inconsistent, and may miss subtle issues until significant scrap or customer returns occur [2]. There is a growing need for intelligent, automated systems that can monitor production in real time, ensure process consistency, and predict failures before they happen.

1.1 Holistic Manufacturing Intelligence and Yield Challenges

The concept of *Manufacturing Intelligence for Reliability and Automated Insights (MIRAI)* refers to an integrated data-driven approach that looks at the entire manufacturing process right from assembly stations on the factory floor to field performance in order to extract insights that improve yield. A holistic approach is necessary because yield loss can stem from many sources: machine settings, component variability, environmental factors, human errors in following SOPs, *etc*. [3] Conventional yield analysis in high-volume electronics production involves pouring over vast amounts of test data and process logs to find correlations with failures [4]. For example, if a particular assembly station shows a drop in yield (more units failing its test), engineers must determine if a systemic cause exists, such as a mis calibrated tool or a change in a component lot. Historically, such analysis has been reactive and limited by human bandwidth and experience [5]. Intel's manufacturing group reported that manual end-of-line yield analysis was too slow and could not examine every unit, prompting a shift to *proactive*, AI-driven yield analysis that examines 100% of data and flags issues for engineers [5]. Advanced analytics, including machine learning, enable detection of yield patterns and root causes more quickly than traditional methods, allowing engineers to address problems sooner and thus prevent further yield loss [5]. This aligns with the broader Industry 4.0 trend of employing big data and AI to achieve adaptive, automated process control for better quality [6].

1.2 Standard Operating Procedures and Human Factors

At the same time, a significant portion of quality issues in assembly processes can arise from human operators not perfectly following the SOPs. SOPs are detailed instructions that standardize how each manufacturing step should be performed to ensure consistency, safety, and quality [7]. Despite training, operators may inadvertently skip steps, use incorrect techniques, or deviate due to fatigue or time pressure [8]. Historically, ensuring Process observation has been done via periodic manual observation by line supervisors or quality engineers. However, manual monitoring is inherently limited, since a supervisor might spot-check an operator for a few minutes, providing only a snapshot of compliance. It's been observed that such manual audits are inconsistent and infrequent, varying by who does the checking and often failing to catch issues that occur between audits [2]. Moreover, compiling and analyzing handwritten compliance notes from different shifts or lines can take days, delaying any corrective actions [2]. Human monitoring of SOP compliance does not scale well to a large production floor and may miss trends that impact yield.

Recent advances in computer vision (CV) and AI offer a promising solution: automated, continuous monitoring of Process observation using cameras and machine learning. By training vision models to recognize the required actions or steps in a process, manufacturers can get an objective, 24/7 "eye" on every station [2]. Unlike a human who can only watch one station periodically, an AI-driven vision system can concurrently monitor all stations and detect any deviation in real time [2]. For instance, a deep learning model can be trained to detect whether an operator picks up the correct component, uses the proper tool, and performs assembly steps in the correct sequence [7].If a step is skipped or done out of order, the system can instantly flag it [7]. This level of constant vigilance ensures that mistakes are caught immediately, allowing for quick correction before a faulty unit moves further down the line [2]. A case study at a display assembly line noted that deploying AI smart cameras for SOP monitoring allowed them to "keep an eye on every movement on the assembly line all the time," something previously impossible with manual checks [2].The result was more consistent adherence to process and the ability to identify operators or steps that needed improvement in real time. Indeed, computer vision-based SOP compliance systems eliminate human bias and fatigue from the equation, providing consistent and objective monitoring around the clock [8]. This not only prevents defects due to process deviation but also creates a rich data source on how the process is actually being executed.

1.3 Data-Driven Failure Prediction

Beyond immediate yield improvement, there is a strategic opportunity in correlating process data (including SOP deviations) with downstream failures to predict and prevent those failures. In electronics manufacturing, some defects are not detected until later stages of production or even until products are in the field (customer usage) [9]. By then, the cost of the failure is much higher which may require scrapping an assembled unit, performing costly rework, or handling warranty returns. If we can predict which units are likely to fail final tests or in the field based on early indicators, we can intervene proactively. Prior research and industry practices have shown the value of linking manufacturing process data to failure outcomes: for example, IntraStage (a manufacturing analytics provider) demonstrated that by correlating detailed test data from production with the results of failure analysis on returned units, manufacturers could identify patterns (signatures) that reliably indicate a unit with a latent problem [10]. Once those "attributes of failure" are known, a predictive engine can scan all in-process and shipped units to find others with the same risk factors and thus target them for preventive action [10]. In essence, if certain process deviations or test parameter anomalies are found to strongly correlate with later failures, they become features in a failure prediction model.

With sufficient historical data, a machine learning model can be trained to recognize the combination of signals that foretell a likely failure (either at end-of-line testing or in field use) [1]. Such a model can then run in real time during manufacturing, alerting operators or stopping a line when a high-risk unit is identified, so that the unit can be inspected or fixed immediately. This approach moves quality control from detection to prediction, saving time and cost by addressing issues before they fully manifest [11]. It also reduces work-in-progress (WIP) waste – rather than adding value to a unit that will eventually be scrapped, the process can be halted or corrected early on.

1.4 Research Objective

In this paper, we propose an integrated system that combines MIRAI workflows for data-driven analysis with computer vision-based SOP adherence monitoring through *Process Observation and Statistical Prediction for Enhanced Compliance and Throughput (PROSPECT)* to improve yield and enable failure prediction in electronics manufacturing. Unlike siloed solutions that either focus on big data analytics or on vision inspection, our approach unifies these into a closed-loop intelligence system. By doing so, we aim to achieve two main outcomes: Improved Yield and Quality, through rapid identification of yield detractors (whether they be machine anomalies or SOP issues) and enforcement of process discipline; and (2) Proactive Failure Prevention, through correlation of process deviations with failure modes and real-time prediction of failures to enable intervention. We emphasize a workflow that not only analyzes historical data (for root cause analysis) but also actively monitors and controls current production (for immediate corrections). The work described is generic and can be applied to many electronics manufacturing settings, focusing on data and algorithms rather than any proprietary equipment. We also align our approach with comparable strategies reported in literature and industry. For example, the use of computer vision for real-time quality and compliance monitoring is increasingly recognized as a transformative technology in manufacturing, with the market for such solutions projected to reach $39 billion by 2029 [12]. Similarly, manufacturing case studies have found that when SOP compliance meets expectations, line efficiency and yield are maximized [13]. Building on these insights, our contribution is to design and document a comprehensive framework that integrates these elements (data analytics, vision, and machine learning) and to discuss its implementation details and results.

2 Methodology

2.1 MIRAI Data Intelligence Workflow

The **Manufacturing Intelligence for Reliability and Automated Insights (MIRAI)** workflow is a data engineering and analytics pipeline designed to empower engineers with self-service insights across the manufacturing process. MIRAI aggregates production data such as test results, process parameters, component information and provides various analysis modules to pinpoint issues affecting yield and reliability. The MIRAI workflow comprises four main components:

2.1.1 Yield Analysis

A self-service analysis tool for identifying signals associated with station yield drops during both ramp-up and sustaining production phases. In ramp-up (new product introduction or early production), yields can fluctuate as the process stabilizes; in sustaining (high-volume steady production), any sudden yield drop is a concern [14]. The MIRAI yield module continuously monitors yield at each test station (the percentage of units passing at that station) and triggers analysis when a significant drop or negative trend is detected. It automatically searches for *common factors* among the failing units that could explain the yield loss. For example, if Station 5's yield fell from 98% to 92%, MIRAI might analyze dozens of attributes of each unit (such as which assembly line it came from, who the operator was, which lot of components were used, calibration settings of the equipment, *etc.*) to find statistical correlations. Techniques akin to commonality analysis are employed such as using association rule mining or contingency table analysis to find factors that are overrepresented in failed units [15]. This helps identify systematic causes of yield loss (as opposed to random defects). The result of a yield analysis might reveal, for instance, that "Units that failed at Station 5 are 3 $\times$ more likely to have come from Line 2 and used Component Batch X," pointing engineers to investigate Line 2 or that component batch. These analytics are delivered in a self-serve dashboard, allowing process engineers to drill down without requiring data science expertise. By quickly pinpointing likely causes, the team can take corrective actions (machine maintenance, station recalibration, additional operator training, *etc.*) to bring yield back up. This workflow reduces the time to root cause by automating much of the heavy data analysis that engineers would otherwise do manually.

2.1.2 Field Failure Analysis

A self-serve analysis capability focused on identifying signals in manufacturing data that correlate with field failures (*i.e.* units that pass all factory tests but later fail during use by customers, resulting in returns or repairs). Data for this analysis comes from linking field return records (or failure analysis reports from returned units) with the original manufacturing data of those units. The MIRAI field analysis module takes a population of products – some that had field failures and some that did not – and performs retrospective data mining to find what factors in the production or test data are predictive of those failures. For instance, it may analyze whether certain test measurements were marginal (close to spec limits) on units that eventually failed, or if a particular factory, production date, or supplier lot is statistically associated with higher field failure rates. This is essentially a supervised learning or statistical correlation task: the units have a label of "field failure" or "no failure," and the system examines all available manufacturing attributes to see which correlate strongly with the label [16]. As with yield analysis, commonality or classification techniques are applied, but with field failures as the target. The outcome could be a set of risk factors – *e.g.*, "Units that failed in the field tend to have had longer soldering cycle times on average" or "Field failures are concentrated in products built with PCB supplier Y in a given week." These insights allow engineering and reliability teams to initiate corrective actions such as design modifications, supplier changes, or targeted recalls for suspect lots. By making field-failure analysis self-service, MIRAI enables a faster feedback loop from customer experience back to manufacturing. This is critical in avoiding widespread issues; as soon as a pattern is detected linking field issues to a process variable, that information can be used to improve production or screening tests [17]. Prior industry approaches that correlate manufacturing test data with field outcomes have shown the value of such analysis in preventing future failures [10] that effectively turns large datasets of past production into actionable knowledge to improve product reliability.

2.1.3 "Bring Your Own Data" Analysis

In addition to standard yield and field analyses, MIRAI supports user-requested custom analyses, essentially allowing engineers to bring their own data for specialized investigations. Often in manufacturing, engineers run experiments or additional stress tests (for example, ongoing reliability tests like ORT, Highly Accelerated Life Testing (HALT), or qualification tests on samples) and want to analyze the results in context of manufacturing data [18]. In the MIRAI workflow, a user can provide a list of units and a binary outcome (Pass/Fail) from some external test or criteria – for instance, a set of units that underwent an On-going Reliability Test (ORT) where a few units failed while others passed. The DataOps team (or the data platform automatically) will gather all relevant manufacturing data for that population of units and perform analysis similar to the above to find signals differentiating the fail group from the pass group. This could involve feeding the combined dataset into a classification model or running targeted queries (like comparing means of certain measurements or doing decision tree analysis to find splits). The result is a report to the requesting engineer with any statistically significant factors that correlate with the failures. For example, if out of 100 units tested in ORT, 5 failed, the analysis might find that all 5 failing units were processed on a particular SMT (surface mount technology) line or all used a particular lot of a component, suggesting a latent issue. By allowing ad-hoc data analysis in this manner, MIRAI becomes a flexible analytics extension for quality engineering experiments. It essentially leverages the data pipeline to answer one-off questions. The pass/fail labeled dataset provided by the user might come from reliability tests (like ORT, environmental stress screening), customer returns classified by failure mode, or even simulations [19]. MIRAI treats this like a mini "competition" between variables to explain the outcome, employing anything from logistic regression to more advanced feature importance ranking. Importantly, this step often requires careful data integration – ensuring that the units in the provided list are correctly matched to their records in various manufacturing databases (traceability, test results, repair logs, *etc*.). The DataOps aspect implies that data engineers may assist in data cleaning and preparation, but the goal is eventually to make this process streamlined so that an engineer with minimal coding can get results by simply uploading a CSV of serial numbers with labels. This workflow dramatically accelerates root cause analysis for issues discovered outside the standard test flow, by bringing all available production data to bear on the problem.

2.1.4 MIRAI Sentinel (MIRAI Sentinel)

The final component is a proactive auto-analysis and alerting system that continuously scans manufacturing data across all build stages to catch emerging issues without waiting for human requests. While the previous components are triggered by an engineer's query or an obvious yield drop, *MIRAI Sentinel* is an always-on watchdog. It performs automated commonality analysis on recent production data in search of anomalies or deviations from baseline. For example, it may automatically cluster recent failures across different stations and see if they share any common factor (*e.g.*, all from the same shift or same supplier lot) and then alert engineers that "5 failures have occurred across two different stations in the last day, all involving Component

Z – this is unusual and worth investigating.” Likewise, MIRAI Sentinel can be configured to monitor trends such as gradually declining yields, shifts in test measurement distributions, or increasing retest rates. When certain thresholds or abnormal patterns are detected, the system triggers an alert or generates an “Auto Commonality Report.” This proactive analysis uses a combination of statistical process control (SPC) rules and machine learning anomaly detection. It might leverage control charts for yield and test metrics and apply clustering algorithms to group suspect units [20]. By spanning across build stages, it means MIRAI Sentinel can connect the dots (for instance, noticing that a particular assembly issue in an early stage is causing fails only at a later test stage). Alerts could be sent via email or shown on a dashboard, highlighting the suspected common cause. The aim is to shorten the time to discovery of issues that might otherwise only be found after a lot of units have failed. This turns yield management from a reactive “pull” (engineers digging for causes after yield drops) into a proactive “push” model where the system itself highlights potential problems [5]. MIRAI Sentinel therefore acts like an automated quality engineer, continuously learning from data and assisting human engineers by focusing their attention where it’s needed. In practice, implementing MIRAI Sentinel requires robust data engineering: streaming data pipelines, data normalization to compare across shifts and lines, and scalable computing to run analyses frequently (potentially on each new batch or each day’s production data). It also requires careful tuning to avoid false alarms – ensuring alerts are meaningful by using logic to filter out spurious correlations (as commonality analyses can sometimes find coincidental patterns [15]). Therefore, MIRAI Sentinel adds an intelligence layer on top of the manufacturing process that preemptively detects and communicates issues, thus embodying the notion of *holistic intelligence* by looking broadly and acting in real time.

From a system architecture perspective, MIRAI is built on a centralized manufacturing data lake that ingests data from various sources: automated test equipment outputs, production execution systems (with information on lots, machines, and operators), as well as external data like field returns [21]. A key enabler for MIRAI’s effectiveness is this integration of data sources. Recent reports from industry emphasize that connecting all plants and processes via an operational data lake to get a real-time, unified view is a foundational step for deploying AI/ML interventions for yield improvement [22]. Our implementation follows this principle – all relevant data about each unit (its genealogy through the factory, all test readings, and eventually the SOP deviation data from vision systems described later) are linked via a unique identifier (such as the unit’s serial number). This comprehensive data foundation allows the analyses in MIRAI’s four workflows to be performed accurately and consistently. The user interfaces for MIRAI include dashboards for engineers with interactive filters and visualizations, and a query engine for advanced users to run custom queries or machine learning models. In essence, MIRAI serves as the analytical “brain” of the manufacturing line, ingesting raw data and outputting insights or alerts that drive improvement actions.

2.2 Computer Vision System for SOP Monitoring

An integral part of our approach is the use of computer vision to monitor station activities for Process observation. This system provides the *eyes* on the factory floor to complement MIRAI’s data analytics [23]. The computer vision setup consists of cameras installed at critical operator workstations (assembly or test stations where human interaction is involved) and an AI inference pipeline that processes the video feed from these cameras in real time. The goal is to automatically verify whether each operator is following the prescribed steps in the SOP for that station and to record any deviation or departure from the SOP.

2.2.1 Camera Installation and Data Capture

Cameras are positioned to get a clear view of the workspace and the operator’s actions, without obstructing the operation. We used industrial-grade cameras with appropriate resolution and frame rate to capture necessary details for example, identifying tools, parts, and hand movements [24]. In some cases, a single wide-angle camera per station is sufficient; in others, multiple angles or a depth camera might be employed if the task is complex. The system design can accommodate both edge processing (smart cameras with onboard AI accelerators like NVIDIA Jetson devices [2]) or a central server approach where video is streamed to a local server running the models. In our pilot, to minimize network load, we opted for edge AI cameras that perform on-board inference and send only summary data/events to the central database. Each camera is time-synchronized with the production line system and tied to a specific station ID. Through the line control system, we know which product serial number is at that station at a given time (since each unit is scanned or otherwise identified at station entry). This integration is crucial: it allows us to tag any detected SOP deviation with the specific unit (serial number) and step, feeding that information into the manufacturing data records.

2.2.2 Model Training for Action Recognition

Developing the computer vision model requires training it to recognize the key actions or objects involved in the station's SOP. This is formulated as an action recognition or sequence verification problem. We collected training data by recording many instances of the station operation, including both correct procedures and some examples of incorrect actions (if available). Depending on the use case, different AI techniques can be used:

(1) Object Detection and Pose Estimation: For tasks where the SOP involves using certain tools or parts, object detection models (*e.g.*, based on convolutional neural networks like YOLO or Faster R-CNN) are trained to detect the presence and placement of those tools/parts in each frame. Human pose estimation models can track the operator's hands and body to see if they reach the correct areas in the correct order. For example, if SOP says "pick up screw, use screwdriver on location A, then B," the system would detect the screwdriver and the motion of hand to location A then B.

(2) Action Sequence Modeling: In more complex workflows, we use sequence models. A common approach is to break the video into a series of discrete actions using a temporal action segmentation model. Alternatively, treat it as a classification per time window: *e.g.*, a deep learning model (such as a 3D CNN or a transformer-based video model) that can classify what action is being done in a short clip. We trained such models on annotated video: subject-matter experts labelled a number of video clips with the action being performed (or labeled if a step was done wrong). The model learns to discriminate correct vs incorrect actions.

(3) State Machine with Vision Triggers: In some implementations, it is useful to encode the expected order of operations as a state machine or rule-based logic, and use the vision algorithms to confirm each step. For instance, state 1 "tool X picked up" must occur before state 2 "tool X applied to part Y". The vision system outputs events like "tool X detected in hand" or "part Y present in fixture" which are fed into a simple logic engine that verifies the sequence.

For our pilot, we started with a relatively constrained task (a single station with a well-defined set of steps) and trained a deep learning model to detect a few key events: whether the operator performed a required check with a camera (vision inspection step) and whether a certain component was tightened with a torque tool. The model was a custom CNN that took image frames as input and output whether the specific action was observed. We augmented this with sensors data when available (*e.g.*, the torque tool provides a reading when used – which we also log for cross-reference). All AI models were developed using open-source frameworks and we ensured not to hard-code any proprietary features. They were validated to a high accuracy on a test dataset of annotated videos before deployment (achieving, for example, $>95\%$ precision and recall in detecting the presence or absence of the critical action).

2.2.3 Real-Time SOP Compliance Monitoring

Once deployed, the vision system operates continuously during production. Real-time inference on the camera feed compares the ongoing operator actions against the SOP model. If every expected step is observed in the correct order, the system remains silent (or just logs compliance). If a deviation is detected – for example, a step is missed within the allotted time or an incorrect action is performed – the system immediately raises an alert. In our implementation, the alert is both visual (displayed on a dashboard for line supervisors) and logged electronically. The alert includes details like: Station ID, timestamp, description of deviation (*e.g.*, "Step 3 – connector inspection – was skipped"), and the unit's serial number. At that moment, a supervisor can intervene, or the system could even be configured to stop the conveyor/belt for that unit if automatic interruption is desired (in our pilot, we opted to alert rather than stop, to study the occurrences first). This immediate feedback mechanism prevents the unit from silently continuing down the line with an undetected process defect. It also provides an opportunity to correct the mistake: the operator or a rework technician can perform the missed step or verify the product before it moves on. Such real-time alerts greatly reduce the chance of a defective unit reaching the end of the line or, worse, the customer [2]. Moreover, continuous monitoring generates a trove of compliance data. The system essentially produces a timestamped event stream of all deviations (and potentially confirmations of correct steps). This data is invaluable for analysis – for instance, to see if certain times of day or certain operators have more deviations, or which steps are most problematic.

2.2.4 Data Logging and Integration

All detected deviations (and optionally a record of compliance events) are stored in a database, with references to product serial numbers and step identifiers. We structured a Deviation Log that captures: (Unit Serial, Station, Step/Action ID, Deviation Type, Timestamp, Operator ID (if available)). Alongside, the production system provides the information of whether the unit eventually passed or failed subsequent tests, was repaired, and so on. By integrating this log with the main manufacturing data (as part of MIRAI's data lake), we can perform correlation analysis between SOP deviations and yield or failures – this is the core of Phase 2 and Phase 3 of the pilot, described next. It is worth noting that careful attention was paid to time synchronization and data alignment. We used the station's start trigger (when a unit arrives and is scanned) to mark the beginning of an operation, and we buffered any vision-detected events during that operation to associate with that unit's serial. This ensures the deviation data is properly linked to the correct unit, which is critical for accurate analysis. Privacy and worker acceptance were also considered: the purpose of cameras is to improve the process and training, not to surveil workers punitively. We ensured the system focused on task elements (and the footage was not used beyond the scope of process improvement), which helped in gaining cooperation for the pilot.

2.3 Process Observation and Statistical Prediction for Enhanced Compliance and Throughput (PROSPECT) Workflow

With the vision system in place to capture SOP deviations, we designed a pilot study in three phases to leverage this data for yield improvement and failure prediction.

2.3.1 Phase 1: Monitor PROSPECT and Identify Key Deviations

The first phase focused on establishing baseline SOP compliance levels and determining whether non-compliance was contributing to yield loss at the station of interest. We selected a particular assembly station that had experienced periodic yield fallout (lower first-pass yield) in the past, suspecting operator errors as a possible cause. Initially, we measured the station's yield fallout rate (the fraction of units failing at that station) over several production runs to have a baseline. Next, we activated the computer vision monitoring at this station to record station activities continuously. Over a period of several weeks, every action at this station was observed by the AI system as described earlier. During this time, we did not make major interventions; the idea was to passively collect data on how often and what types of SOP deviations were happening. The system generated a log of deviations, which we then analyzed. We identified key SOP deviations by frequency and potential impact. For example, we discovered that one particular step, scanning a barcode on a sub-component to verify its presence, was occasionally skipped. Another deviation noted was an improper torqueing sequence: operators sometimes tightened screws in the wrong order or missed the last screw, contrary to the SOP. We also noted the frequency of each deviation and whether certain operators had more deviations, though individual performance was anonymized in analysis. This phase had an *iterative loop* aspect: when a critical deviation was identified, we took immediate corrective action by communicating with the production team. For instance, upon finding the skipped barcode scans, we updated the station's work instructions and retrained operators to emphasize that step. We also added a simple error-proofing measure: the station software now requires the barcode scan input before allowing the process to continue (forcing compliance). These interventions (training and process changes) were implemented, and the station yield was measured again to see if it improved. Indeed, after addressing the top deviations, the station's yield fallout dropped noticeably (we observed an improvement from about 92% first-pass yield to 96%, for example, after enforcing the barcode scan step). Phase 1 is thus a cycle of *observe* → *identify* → *fix* → *observe again*, gradually reducing human error-induced falls in yield. In essence, this phase answers: "What SOP violations happen and are they hurting yield?" It establishes a direct link between adherence and quality, echoing the industry observation that SOP compliance correlates with better performance [25]. By the end of Phase 1, we had a much cleaner process at the pilot station (fewer deviations after interventions) and a list of residual deviations that were harder to eliminate or quantify. Crucially, we had captured data that some deviations still occurred (albeit less frequently), and those instances could be studied in Phase 2 for their impact on failures.

2.3.2 Phase 2: Track Deviations by Serial Number and Correlate with Failures

In this phase, we shifted from focusing on station yield at the point of occurrence to the *downstream effects* of SOP deviations. The approach was to follow each unit through the rest of the manufacturing process (and even field use, if data allowed) and see if those that experienced a deviation at the station have a higher chance of failing later compared to those with no deviations. We started by instrumenting the system to track serial numbers of units with deviations. For every unit that passed through the monitored station, the deviation log was checked. If any SOP deviation was recorded for that unit, we flagged that unit in a "deviation present" category; units with no detected issue were flagged as "deviation-free." We then compiled the outcomes for each unit: did it pass final testing? Did it require rework or repair? If it failed, what was the failure mode (captured via failure analysis or troubleshooting logs)? If available, we also tracked if the unit had any field return or early life failure after shipment. This data was gathered over many units (on the order of thousands, to get statistically meaningful results) during the period of the pilot. With this labeled dataset (units with deviation vs without, and their eventual fates), we performed a deviation-failure correlation analysis. Essentially, this is calculating the conditional probabilities and looking for statistically significant differences. For example, we found that units which had the torque sequence deviation (missed screw tightening) were far more likely to fail the end-of-line functional test for that product. The failure mode in those cases was often related to that part – *e.g.* a loose heatsink or connector causing a test failure. We quantified this: suppose out of 1000 units that had no deviations, 5 failed later tests (0.5% failure rate), but out of 50 units that had a certain deviation, 5 failed (10% failure rate) – that would strongly indicate a correlation. In our study, one particular deviation (improper torque) had a very high correlation with a specific failure mode observed in environmental stress tests (vibration test failures), with an odds ratio suggesting those units were ~8 times more likely to fail than baseline. On the other hand, some deviations seemed to have little to no impact – *e.g.*, if an operator momentarily deviated but corrected themselves (a transient hesitation that was flagged but ultimately the step was done), it did not translate to any measurable difference in outcomes. We also cross-correlated the data: it could be that a combination of deviations or a deviation at one station in combination with another factor leads to failure. However, since our pilot dealt with one station primarily, we kept the analysis straightforward: a binary "deviation happened at station X" vs outcomes. The failure analysis (FA) data from the repair technicians was invaluable – it allowed us to link a cause to effect (for example, "unit failed final test due to loose connector; indeed, a deviation earlier indicated that connector was not scanned or secured properly"). We measured the deviation-failure correlation in terms of metrics like precision and recall as well: if we use "deviation occurred" as a predictor of failure, how accurate is it? For critical deviations, the precision (how many of the flagged units actually failed) might not be extremely high because many units with a deviation still pass (perhaps the deviation was minor or caught later), but the recall (how many of the failing units had a known deviation) was quite high. In one case, 70% of the units that failed a certain test had experienced a particular SOP deviation upstream. This kind of insight validates the hypothesis that PROSPECT has a direct effect on yield and reliability. It also provides a list of deviations ranked by their impact on quality. This information feeds back to Phase 1's loop: deviations that show strong correlation with failures become top priority to eliminate through process improvements or poka-yoke (mistake-proofing) mechanisms. Essentially, by the end of Phase 2, we had created a deviation-failure repository – a collection of cases linking specific procedural missteps to specific failures, complete with data statistics. This repository is an asset for both engineering and training: it can be used to justify investments in automation or training (*e.g.*, "We must fix this step because it's causing X% of our failures") and to educate operators on the importance of each SOP step ("Skipping this screw tightening leads to failures in vibration testing, as data shows"). Moreover, this set of correlated features and outcomes lays the groundwork for predictive modeling.

2.3.3 Phase 3: Build and Train a Failure Prediction Model

In the final phase, we leveraged the insights and data collected to develop a machine learning model that predicts unit failures in real time based on observed SOP deviations (and potentially other data). The concept is to enable the factory to catch a likely-failing unit as early as possible and apply a fix or additional screening right away, thereby preventing the failure from either propagating down the line or escaping to the field. The input features to the model included the SOP deviation flags for each unit from the monitored station (and we can extend to multiple stations as we scale up). For our pilot, since we instrumented one station, the primary features were binary indicators of whether each type of deviation occurred for that unit. We also considered adding other easily available features to improve prediction – for example, whether the unit had any borderline test results (within spec but near limit) at that station, or how many times the unit was retested at that station. But the simplest effective model was one that used the presence/absence of the key deviations as features. The target label for the model was whether the unit eventually failed at final test (or required any repair) – essentially a proxy for yield outcome. (In future extensions, the target could be field failure, but that data was scarcer; for the pilot we focused on predicting final test fallout, which itself is highly beneficial for yield if addressed). We split our collected dataset (Phase 2 data) into training and validation sets, maintaining chronological order to avoid leakage (training on earlier units, and validating on later units, mimicking deployment). We then trained a classification model. We experimented with a few algorithms: a simple logistic regression, a decision tree, and an ensemble like a random forest or gradient-boosted trees. Given the relatively small number of features and their categorical/binary nature, even logistic regression was quite interpretable and effective – it gave weight to each deviation type corresponding to how predictive it was. The ensemble models gave a slight performance boost by capturing interactions (for instance, if two different deviations together made failure even more likely). Ultimately, we chose a gradient boosted decision tree model (similar to XGBoost) for deployment, as it handled feature interactions well and provided good accuracy without overfitting. The model was trained to output a probability that a unit will fail, given the observed deviations. On the validation dataset, we achieved an accuracy in the range of ~90% for predicting failure vs pass, with a high recall for failures – meaning it caught most of the failing units (for example, ~85% of the units that did fail were assigned high risk by the model). We tuned the threshold of the model to favor capturing failures (even if it meant some false positives), because the cost of a false positive (some extra inspection) is much lower than the cost of a false negative (a bad unit slipping through). In practice, one can adjust this threshold based on business needs (*e.g.*, how much re-inspection capacity is available).

After training and offline validation, we moved to deployment of the failure prediction model. We integrated the model into the station's software such that after processing each unit (or at the end of the line, before final test), the system would automatically evaluate: if the unit had any SOP deviations logged, it feeds those into the model (a simple lookup and calculation) and produces a risk score. If the risk score exceeds a predetermined threshold, the system flags that unit for immediate attention. During deployment, this meant the unit was routed to a special inspection station before final testing. At that station, a technician would double-check the unit for the likely issue (for example, if the model flags "high risk due to missed screw tightening", the technician will specifically check all screws and perform the missed step). In many cases, this predictive interception allowed us to fix the problem such that the unit then passed final test, improving the first-pass yield. If the unit was flagged but nothing obvious was found, we still ran it through all tests and kept it under observation (none of the flagged units were sent to customers without thorough vetting). Over time, we measured the effectiveness: the number of units that would have failed final test but were fixed due to early prediction. This is essentially the true positive count of the model. We also tracked the false positive rate (units flagged that would have passed anyway) to ensure it was at a manageable level. The model's performance was very encouraging – for instance, in a month of operation, out of the units the model flagged, a significant portion indeed had issues that required rework (caught early instead of later), and the overall end-of-line yield improved by a few percentage points as a result of these pre-emptive fixes. This aligns with the goals set out: predict and fix potential failures in real time, thereby increasing the first-pass yield and reducing waste. In broader context, the predictive model effectively extends the reach of our quality control: instead of relying purely on final test outcomes, it uses process deviations as predictive signals. It is a form of predictive quality analytics that shifts us from "find and reject bad units" to "anticipate and correct bad units" – a hallmark of advanced smart manufacturing systems [1].

It's worth noting that as we accumulate more data (Phase 3 is ongoing in a sense), the model can be retrained and improved. If additional stations are instrumented with vision systems, their deviations can be added to the feature set, making predictions even more comprehensive. The modular nature of the system means we can plug in more data sources (*e.g.*, machine sensor data or operator biometric data) if they prove predictive. But even with just SOP deviation data from one station, we demonstrated a clear value: a measurable improvement in yield and a reduction in escaped defects. The deployment also provided real-time feedback to operators – knowing that deviations immediately trigger scrutiny created a positive pressure to follow SOPs more rigorously (this was anecdotal but observed). Phase 3 closes the loop by enabling real-time intervention: the moment a risky situation is detected (either by direct deviation alert or by predictive flag), action is taken to either correct the process or isolate the unit for repair. This embodies the synergy of integrating MIRAI data analysis with computer vision: we not only analyze and understand problems but also actively prevent them on the line.

3 Results

We evaluated the integrated MIRAI and PROSPECT system through a pilot deployment in an electronics manufacturing line. The results are presented in two parts: (1) insights and improvements gained from the MIRAI analytics and SOP monitoring (Phase 1 and Phase 2 outcomes), and (2) performance of the failure prediction model and its impact on yield (Phase 3 outcomes). All results are reported in a generic context (no proprietary data) but reflect the scale of a real manufacturing scenario.

3.1 Yield Improvement and Process Insights

3.1.1 Station Yield Recovery

In Phase 1 of the PROSPECT, after implementing continuous SOP monitoring and subsequent interventions, the target station's yield showed notable improvement. Initially, the station's first-pass yield (FPY) was fluctuating and averaged around 92% (meaning 8% of units required rework or failed at that station). By identifying the most frequent SOP deviations (such as missed scans and incorrect torque sequence) and addressing them through operator retraining and process enforcement, we observed the FPY rise to ~96% over the following production cycles. This ~4 percentage point improvement is significant in a high-volume environment, representing dozens of units per week that no longer needed rework. It directly translates to cost savings and increased throughput. More broadly, across the pilot period, the overall line FPY (cumulative yield through all stations) also improved, although the pilot only focused on one station's changes. This suggests that fixing issues at one station prevented a cascade of problems down the line. These findings reinforce the often-stated manufacturing principle that adherence to "One Best Way" procedures yields better performance [25]. In fact, our data provided a quantitative example of that – when SOP compliance approached 100% for the critical steps, the station efficiency and yield were at their highest. This result mirrors other industrial case studies where plants that achieved high SOP compliance saw corresponding high line performance [26].

3.1.2 Deviations Frequency and Reduction

Over the course of Phase 1 and Phase 2, we tracked the frequency of SOP deviations at the station. Initially, in the first two weeks of monitoring, deviations were detected in roughly 15% of the units processed (some minor, some critical). After feedback and corrective measures were introduced (*e.g.*, making a barcode scan mandatory, reinforcing training), the deviation rate dropped to under 5% of units. This demonstrates the effect of simply measuring and responding, operators and supervisors became aware that certain mistakes were happening and took steps to avoid them. Among the types of deviations, we found that procedural misses (completely skipped steps) were less frequent but often more impactful, whereas sequence or timing deviations (steps done out of order or too quickly without verification) were more common but sometimes had less impact if eventually corrected. By the end of the pilot, the most egregious deviation (the missed scan) was virtually eliminated, while a few others (like slightly out-of-order operations that did not affect the outcome) still occurred occasionally. The comprehensive monitoring made it possible to sustain this improvement; unlike a one-time audit, the AI system continuously ensures that the process does not drift back to old habits. From a management perspective, this data allowed us to pinpoint where additional training was needed – for instance, if one shift had more deviations than another, management could investigate why (perhaps a less experienced operator on that shift, *etc*.) and take action.

3.1.3 Correlation of Deviations with Failures

In Phase 2, our analysis provided concrete evidence linking SOP deviations to downstream failures. One striking result was the correlation between the missed torque step and a failure in a subsequent vibration test (part of reliability testing). Out of all units that had the torque deviation, 20% later failed the vibration test (due to things like loose components), whereas among units with no such deviation, only ~2% failed the same test. This tenfold difference strongly indicates causation – improper torque likely caused components to be insufficiently secured, which then led to failures under vibration stress. When presented with these findings, the manufacturing engineers were convinced to implement additional safeguards (they decided to introduce a sensor to verify torque for each screw, adding an automated check in addition to the vision). In another example, a skipped inspection step correlated with an increase in cosmetic defects seen at final quality check. While those cosmetic issues didn't cause functional failures, they did result in rework (polishing or reassembling parts), impacting efficiency. Units that skipped the inspection had a 15% cosmetic rework rate versus 5% normally. By correlating each deviation type with various outcome metrics (final test fails, reliability fails, rework incidents, and even warranty returns for the period we could observe), we built a matrix of influence. This kind of data is rarely available in traditional operations, as the links are not traced. But here we had a clear mapping: for each SOP deviation type, we could quantify its effect on yield or quality metrics. The repository indicated, for example:

(1) Deviation A (missed step): associated with failure mode X, correlation strength: high.

(2) Deviation B (incorrect sequence): mild correlation with extended test time, but no direct failures (operators usually caught up and corrected later).

(3) Deviation C (skipped verification): moderate correlation with field returns of issue Y,

suggesting a latent defect might slip through.

These insights not only validated the approach but also gave direction for continuous improvement. They essentially told us where to focus engineering effort. Additionally, from a Six Sigma perspective, we considered the deviations as a source of process variation. By eliminating those deviations, we reduce variability in the process, which naturally improves yield (higher sigma level). Our results empirically demonstrate this: the variance in yield results at the station narrowed after Phase 1, and the overall defect rate decreased after addressing the high-impact deviations identified in Phase 2.

3.1.4 MIRAI Analytics Outcomes

Concurrent with the SOP pilot, the MIRAI platform's yield and field analysis modules were run regularly on the production data. While the MIRAI system covers the entire line, for brevity we note a few key outcomes that intersected with our pilot:

(1) The MIRAI yield analysis module independently flagged the pilot station for having an unusual uptick in failures during the initial baseline period, correctly identifying that most fails were associated with a specific operator and shift (which corresponded to the time the missed scan issue was occurring frequently). This was a good cross-validation; MIRAI's automatic data crunching pointed to a human factor issue at that station, which our vision system then directly observed. This shows the synergy: data analytics can highlight "where to look," and vision provides the "what exactly is happening."

(2) The MIRAI field analysis (though based on limited return data in the pilot's timeframe) indicated that units with the vibration failure mentioned above had all been processed at the pilot station by a specific tool ID – again correlating to the torque issue. This kind of finding is inline with industry experiences where manufacturing data patterns are tied to field reliability. It underscores that the impact of SOP deviations can extend to field performance, not just immediate yield.

(3) The BYOD analysis was tested by feeding in some ORT results: a batch of units had undergone an accelerated life test (where a couple failed). MIRAI BYOD analysis found that those failing units were among the ones that had minor process deviations (like shorter solder time) upstream. While not directly part of SOP, it shows the utility of having an analytics pipeline that can incorporate any new data and link it to production info.

(4) MIRAI Sentinel alerts during this period caught a separate issue on another station (unrelated to our main pilot) where yield was dropping due to a misaligned test fixture. This was resolved quickly. We mention this to illustrate that our integrated approach does not rely on only one type of data; the MIRAI system continues to handle machine/equipment issues in parallel, whereas the SOP vision pilot added the human procedure aspect into the holistic view.

Overall, the results demonstrate that integrating these systems provided both *rapid local improvements* (fixing issues at the station) and *broader visibility* into how process execution affects quality. We essentially expanded the feature space of manufacturing data to include human adherence metrics, which proved to be important predictors.

3.2 Failure Prediction Model Performance

3.2.1 Predictive Accuracy

The failure prediction model trained in Phase 3 was evaluated on historical data and then monitored live. On the test dataset of a few thousand units (with known outcomes), the model achieved an AUC (Area Under ROC Curve) of about 0.92, indicating excellent discrimination between units that fail and those that pass. At an operating threshold chosen to prioritize catching failures, the model's sensitivity (true positive rate) was around 85%. This means 85% of units that did end up failing final test were correctly predicted as high-risk by the model *before* the final test occurred. The specificity (true negative rate) was slightly lower, around 80%, since we tolerated some false positives. The precision or positive predictive value was in the range of 30–40%, meaning that among the units flagged as high-risk, roughly a third actually would have failed if not intervened. While 30–40% precision might seem moderate, it is actually quite useful in context – these flagged units can be inspected with relatively low effort, and if 1 in 3 is a true issue, that's a big win considering those would have been failures. In fact, many predictive maintenance or quality models in industry operate in regimes of low base failure rates, so a precision of 30% can be economically justified if the cost of checking a false alarm is small compared to the cost of a miss. We should note that the model was somewhat conservative in that any unit with even a minor critical deviation was flagged. There were almost no false negatives for the specific failure modes related to the monitored deviations; the few failures that slipped through were due to other causes (unrelated to the SOP steps we monitored).

3.2.2 Real-Time Deployment Results

During live deployment over one month, the model flagged approximately 50 units as high-risk out of several hundred produced. Of those 50, about 15 were confirmed to have real issues that would likely have caused test failures or field failures (true positives). These issues were fixed on the spot. For example, one flagged unit was found to have an improperly seated connector (the SOP deviation was a skipped verification step) – the technician reseated it, and the unit then passed all tests. Without the system, that unit would have failed at final test or perhaps passed but failed in the field. Another flagged unit had a missing screw (caught by visual check after flagging) which was then installed, saving that unit from likely failure. The other 35 flagged units (false positives) were re-inspected and no problems were found; nearly all of them passed final test normally. In those cases, the model erred on the side of caution (for instance, an operator might have slightly deviated but corrected it, and the unit was fine, yet it was flagged due to the deviation log entry). We are analyzing those false positives to see if the model can be refined to ignore truly benign deviations (perhaps by incorporating the fact the step was eventually done, albeit late). However, the manufacturing leadership was pleased with this result: 15 units proactively saved from failure is a direct improvement in yield, and the overhead of checking 35 extra units was manageable. In fact, the yield improvement at final test was quantifiable. The line's final test yield improved from ~95% to ~98% during that period. Not all of that is solely due to the model (some general improvements happened too), but a portion can be attributed to catching those failures early. Even more importantly, every unit that is fixed early saves significant time; a unit caught at the station can be reworked in minutes, whereas if it fails at final test, it disrupts the flow and requires sending the unit to a repair area, retesting after fix, *etc*., which could take hours. So there is an efficiency gain beyond the yield percentage.

3.2.3 Case Study – Preventing a Field Escape

While our deployment time was short to gather field data, one notable anecdote stands out. One unit was flagged by the model for a minor SOP deviation (the operator did not follow the exact order of two sub-steps, but eventually completed them). The unit passed final functional tests, so normally it would have shipped. Because it was flagged, the quality engineer decided to put it through an extra stress test overnight. It turned out that under prolonged thermal cycling, the unit did fail due to a joint that was not perfectly soldered (the deviation might have caused a suboptimal solder reflow). This unit was caught and scrapped before shipment. While this is a single instance, it exemplifies the potential of failure prediction to prevent a possible field failure (which could have resulted in a costly customer return or warranty claim). It underscores that a predictive model can add a layer of protection especially for latent defects that aren't detectable by normal tests but have telltale signs in the process data.

3.2.4 Integration with MIRAI Sentinel

We also integrated the model's logic into the MIRAI Sentinel platform. Instead of just alerting on correlations, MIRAI Sentinel can use the predictive model to watch all units. In effect, every time a deviation was logged (as part of the data stream), MIRAI Sentinel would evaluate the risk and generate an alert for high-risk unit. This means even if we expand to more stations, a central system can coordinate the flags and possibly even suggest where to route the unit (to an offline check). The result is a unified alert dashboard that not only warns of equipment issues (as it did before) but now also of specific units at risk due to process anomalies. This unified approach is a step towards what some quality experts call a "360-degree view of quality"– combining machine, process, and human factors data to ensure each product meets standards [27]. Our results contribute to that vision by showing how to incorporate PROSPECT data effectively.

3.2.5 Economic Impact

Although this paper focuses on technical results, a brief note on the potential economic impact is warranted. Improving FPY even by a few percentage points on a high-volume electronics line can save hundreds of thousands of dollars annually in labor, scrap, and warranty costs. Our pilot's ~4% station yield improvement and ~3% final yield improvement translate to fewer units needing rework and more units out the door per day. Additionally, preventing field failures avoids not just the direct cost of returns but also intangible costs like customer dissatisfaction. The real-time fix approach also reduces WIP and cycle time, as units do not circulate back and forth for fixes. Thus, the integration of MIRAI and vision we demonstrated has a clear business case, aligning with known benefits of AI in manufacturing such as reduced defects, cost savings, and throughput improvement [28]. Our results are in line with other reports where AI-driven interventions led to yield gains and lower defect rates; for example, an AI-based defect classification system can significantly boost production yield by catching defects early [29]. In our case, instead of optical defect inspection, we caught process defects, but the end goal of yield boost is the same.

3.3 Summary of Key Results

To summarize quantitatively:

(1) SOP deviation rate at target station: reduced from ~15% of units to <5% through Phase 1 actions.

(2) Station first-pass yield: improved from ~92% to ~96% after addressing key deviations.

(3) Correlation example: units with deviation X were ~10× more likely to fail later testing than those without (clearly identifying X as a root cause contributor).

(4) Failure prediction model: 85% of failing units correctly predicted (caught) with ~30% precision in a pilot deployment; final test yield increased ~3% with model in place.

(5) Zero critical failures went unaddressed among those monitored – meaning the combination of vision + model caught all instances of the known issues we targeted.

(6) The system demonstrated scalability in data handling, analyzing thousands of data points (images, events, test records) per unit in an automated fashion.

These results support the hypothesis that a holistic approach combining data analytics, computer vision, and machine learning can substantially improve manufacturing outcomes [30]. They also highlight that neither data analytics nor vision alone would be as effective: it was the combination that allowed identifying and preventing issues. In the next section, we discuss these implications and how they compare to other approaches in the industry.

4 Discussion

The successful pilot implementation of the integrated MIRAI + computer vision + PROSPECT approach provides several insights into both the technical and operational aspects of advanced manufacturing quality systems. In this section, we interpret the results, compare our approach with related work, examine the generalizability of the solution, and discuss challenges and future directions.

4.1 Integration of Diverse Data Sources

One of the standout aspects of our approach is how it brings together traditionally separate data streams – test data and human action data – into one analytical framework. In manufacturing, it's common to have siloed systems: a Manufacturing Execution System (MES) that captures process data and yields, and maybe a separate quality system for audit findings or manual observations. By capturing PROSPECT via computer vision and feeding that into the unified MIRAI data lake, we created a richer dataset for analysis and modeling. This aligns with the Industry 4.0 philosophy of system integration and a "single source of truth" for manufacturing data [22]. Our results show that this integration is not just technically feasible but highly beneficial. For instance, MIRAI's analytics became more powerful when we included the SOP deviation flags as additional features – we could uncover correlations (like the torque issue) that might have been obscured if one only looked at test data in isolation. Comparable approaches in industry often focus on one domain: *e.g.*, automated optical inspection (AOI) systems focus on visual defects on products, and manufacturing intelligence platforms focus on sensor and test data. We effectively combined a "process compliance monitoring system" with a data analytics platform. This holistic view is what gave us a 360-degree understanding of the root causes. Our work operationalized that integration in a custom way, demonstrating that the vision data can feed a predictive model that ties into quality control. A key lesson is that investment in data engineering – to ensure different data modalities can join on common identifiers – pays off greatly. We needed to ensure timing, serial number tracking, and database schema all aligned, which was non-trivial, but once in place, it allowed complex analyses with ease. This suggests future factories should design data architecture with such integration in mind from the ground up.

4.2 Impact on Yield and Quality

The improvement in yield we observed is consistent with the idea that reducing process variation (including variation introduced by human error) improves quality. This echoes fundamental principles of Six Sigma and Lean manufacturing, where standard work and elimination of deviations lead to better outcomes [31]. Our approach provided a high-tech way to enforce and measure standard work. Traditionally, Lean practitioners implement standard work charts and audit them; our system automates that audit and provides quantitative feedback in real time. This can be seen as a form of digital poka-yoke, where the system acts as an error-prevention mechanism by catching mistakes [32]. The yield improvements, while demonstrated at one station, hint at the potential if scaled line-wide or plant-wide. If every critical station is monitored and optimized, incremental improvements at each can compound into a large overall gain (especially in complex assemblies with many steps). Additionally, by catching issues upstream, we reduce the accumulated cost of defects – a defect caught and fixed at station 5 is cheaper than one found at final test or, worse, in the customer's hands. This is in line with the well-known "Rule of 10" in quality (each step later you find a defect, the cost multiplies by roughly 10). We effectively pushed detection as far upstream as possible.

4.3 Comparison with Prior Approaches

It's valuable to compare our integrated approach with other strategies.

4.3.1 Manual SOP Auditing *vs.* Computer Vision

Before vision, companies relied on periodic SOP compliance audits. These are labor-intensive and often too late to prevent defects. As the ADLINK case and our introduction noted, manual monitoring is inconsistent and cannot cover all operations. Our results confirm that an AI vision system can achieve consistent 24/7 monitoring and react in less than a second to issues, something impossible with manual audits. Other researchers and vendors have begun documenting similar successes with vision. This technological shift essentially ensures adherence in ways that were previously only aspirational.

4.3.2 Automated Test Data Analysis

Machine learning applied to test data (without vision) has been used for yield improvement and predictive maintenance. Our MIRAI platform is conceptually similar to those – it uses data to find correlations and root causes. The difference is that we extended the data to include human factors via SOP logs. Many traditional yield analyses might not capture that an assembly step was done incorrectly; they might only see the end symptom (like a measurement out of range). By adding the cause (deviation event) as data, we enhanced the analysis. In essence, our approach could be seen as adding a new category of sensor: the eyes on the process.

4.3.3 Direct Automated Inspection *vs.* SOP Monitoring

One might ask, why not simply rely on direct automated inspections for quality (like vision systems that inspect the product for defects)? Indeed, AOI and end-of-line vision inspection are common in electronics (for solder joint inspection, *etc*.). Those catch defects directly on the product. Our SOP monitoring is complementary: it catches the *process mistake* that might lead to a defect, often before the defect is even visible or testable. This is a proactive *vs.* reactive distinction. Both approaches together would be ideal – inspect the product *and* ensure the process is correct. Ensuring the process prevents many defects from ever occurring, reducing the load on final inspection. This is analogous to how in healthcare, preventing disease (via monitoring and intervention) is better than just diagnosing it later.

4.3.4 Predictive Models in Manufacturing

The use of predictive models (like our failure prediction model) is increasingly common under the umbrella of predictive maintenance and predictive quality. For example, others have applied ML models to predict machine failures or to predict yield of a lot before it finishes processing [33]. Our model specifically predicts product failures based on process deviations. This is somewhat novel because it leverages human error data in the prediction, whereas many predictive maintenance models use sensor data from machines. Our system is like an automated, data-driven FMEA: it identified a cause (SOP deviation) and showed the effect (failure), then we took action to control that cause. The difference is it was based on real data rather than theoretical assessment.

4.3.5 Scalability and Generalizability

While the pilot was on one station, the approach can be scaled to multiple stations and different product lines. MIRAI is inherently scalable as a data platform; adding more stations just means more data, which modern data processing can handle (especially with cloud or on-premise clusters). The computer vision system would need to be replicated for each station type.

This implies training new models for each station's SOP (since each has distinct actions). That is a non-trivial effort, but techniques like transfer learning and more general action recognition models can speed it up. There are also emerging no-code vision platforms that claim to allow quick setup of such monitoring. In an enterprise scenario, one could create a library of vision models for common assembly tasks and deploy them widely. The infrastructure (cameras and compute) cost is a factor, but as vision technology becomes cheaper and more ubiquitous, this becomes more viable. Furthermore, the approach is general to any manufacturing operation where humans perform critical tasks – not just electronics. One could see applications in automotive assembly, medical device manufacturing, or even warehouse operations for quality assurance. The key is identifying processes where deviations significantly impact quality. Our work provides a template: start with a pilot at a pain point, prove the value, then expand. It also shows how to integrate with existing data systems, which is often a concern (people fear new systems that don't talk to old ones). We integrated via the data lake and by aligning with MES events, demonstrating you don't have to rip-and-replace anything; you augment it.

4.3.6 Worker and Organizational Impact

It's worth discussing how this system affects the people on the factory floor. Initially, there can be apprehension that cameras watching operators could be used in a punitive way or create a "Big Brother" environment. We addressed this by focusing the feedback on process, not personal performance, and by involving operators in the improvement process (for example, showing them how eliminating a certain mistake made their job easier by reducing rework). Over time, operators saw the system as a helper – it would catch something they missed, essentially acting as a safety net. Also, with fewer failures, their work actually went more smoothly (fewer angry rework technicians coming back asking about mistakes). Training and communication are vital: we stressed that the goal was to improve the process and help them succeed, not to punish. This approach can actually elevate the role of operators: they become partners in a high-tech process and can take pride in achieving high compliance. In fact, one could gamify PROSPECT (though we did not do this) – showing metrics of improvement and recognizing teams that have zero deviations for a week, *etc*. From an organizational standpoint, this integrated system breaks down barriers between different teams: process engineers, quality engineers, data scientists, and line supervisors all had to collaborate. It fostered a more data-driven culture on the floor. Decisions to change processes were backed by data (*e.g.*, "the data shows this step is causing 80% of our failures, so we will fix it" instead of arguments based on anecdotes). This is an important cultural shift towards what some call *manufacturing intelligence*.

5 Limitations

Despite the successes, there are some limitations and challenges to address.

5.1 Model Scope

Our failure prediction model was limited by the scope of data (one station's deviations). If a failure was caused by something outside that scope (*e.g.*, a PCB defect not related to assembly), the model wouldn't catch it. Thus, it's not a panacea for all failures, only those tied to the monitored parameters. As we scale, we need to include more features to cover more failure modes.

5.2 False Alarms

As seen, there were false positives. Tuning the system to reduce unnecessary alerts without missing true issues is an ongoing effort. This involves both refining vision detection (to not log a deviation unless it's truly a deviation) and refining the predictive model. We might incorporate more context to distinguish a serious deviation from a harmless one.

5.3 Vision Challenges

The computer vision system, while robust for the pilot, can face difficulties in more complex settings. Changes in lighting, obstructions, or operator behavior variations can affect detection. Also, if the product model changes or the process changes, the vision model may need retraining. We discovered that even something like an operator wearing gloves *vs.* not wearing gloves could confuse the model initially (we then included both scenarios in training data). Maintaining and updating these models will require a dedicated effort or a user-friendly training interface. This is a general challenge in AI adoption in manufacturing – the need for updating models as processes evolve.

5.4 Data Volume and Latency

Processing video for many stations could be data intensive. We mitigated this by edge processing (only events go to the server, not full video), but in some contexts storing video might be desirable for later analysis. That raises storage and privacy questions. In our case, we did not need to store raw video long-term, just the detected event timestamps which are tiny in size. So our system is efficient in that sense.

5.5 Generality of SOP Deviations

The types of deviations and their impacts can vary widely by process. In some processes, a deviation might not have any effect (maybe a redundant step). So one must be careful not to overreact to every deviation. Our correlation phase addressed that by quantifying impact. But if someone applied such systems blindly without analysis, they might waste effort on low-impact deviations or, conversely, not realize a critical one. Thus, the combination of automated monitoring with human engineering judgement remains important.

5.6 No Internal Proprietary Tools

We consciously described everything in generic terms. In actual implementation, one might use specific software or platforms (like a specific brand of data historian or a certain AI framework). Our aim was to show the approach without tying it to a vendor. This is beneficial academically because it focuses on principles, but a real company would need to either develop or purchase the specific tools to implement it.

6 Future Work

Building on this pilot, there are several avenues for further development.

6.1 Multi-Station and End-to-End Monitoring

We plan to extend vision monitoring to multiple stations (including automated ones where a robot might perform tasks, to verify the robot did them correctly) and link deviations across the entire process. This could lead to a much more powerful predictive model that uses a sequence of events from multiple stations to predict final quality.

6.2 Advanced AI Models

The action recognition model can be made more sophisticated. For example, using deep sequence models (like an LSTM or transformer taking video frames as input) might capture deviations in subtler ways and reduce false positives. Also, anomaly detection models could be employed so the system can learn what a “normal” operation looks like and flag anything that deviates from the norm, even if not pre-defined.

6.3 Operator Guidance Systems

We are considering integrating augmented reality (AR) or real-time feedback to the operator through an interface. Currently the feedback is mostly an alert to supervisor or a light signal. If operators had, say, AR glasses or a screen highlighting what step to do next or warning them they missed something, it could guide them before a deviation becomes permanent. This would truly close the loop at the operator level – prevention rather than post-fact alert.

6.4 Expansion of MIRAI Analytics

The data collected on PROSPECT could feed other MIRAI modules. For example, a training effectiveness analysis – measuring if after training sessions the deviation rates drop (and how fast). Or feeding into a digital twin of the process that simulates how errors propagate. Also, including cost models in MIRAI to prioritize which issues to fix first based on potential savings.

6.5 Comparative Studies

We intend to benchmark this integrated approach against others. For instance, compare yield improvement purely from a data analytics approach *vs.* with the added vision data, to quantify the incremental benefit of vision. Also, measure ROI in terms of cost of equipment *vs.* savings.

6.6 Generalization to Autonomous Corrections

Ultimately, we envision a system that not only predicts failures but can autonomously correct them or adjust the process. For example, if a deviation is detected, the system might automatically adjust a downstream test to be more stringent for that unit (to ensure the defect is caught). Or if certain deviations keep happening, the system might automatically modify the SOP or machine parameters (with approval workflow) to error-proof it. This would be a step towards a self-optimizing production line.

7 Conclusion

In this paper, we presented a comprehensive approach to improving yield and predicting failures in electronics manufacturing by integrating a Manufacturing Intelligence for Reliability and Automated Insights (MIRAI) system with computer vision-based PROSPECT monitoring. Our solution spans data engineering, real-time monitoring, and machine learning, creating a closed-loop feedback system for process improvement. Through a pilot study, we demonstrated that this integration can effectively identify the root causes of yield loss (including human procedural errors), facilitate timely corrective actions, and enable proactive failure prediction to catch defects before they escape. Key contributions of this work include:

7.1 Holistic Data Integration

We showed how diverse data sources – production test data, operator action logs from vision, and failure analysis results – can be unified and utilized for advanced analytics. This holistic view provided insights that would be inaccessible to siloed analysis, highlighting the importance of integrated manufacturing intelligence in the era of Industry 4.0.

7.2 Computer Vision for SOP Compliance

We implemented a computer vision system to automatically monitor SOP compliance at an assembly station. The system achieved continuous, unbiased observation of operator practices, detecting deviations in real time. By doing so, it effectively digitized the enforcement of standard procedures. Our results confirmed that such a system can drastically reduce human error-related defects, consistent with emerging industry reports of AI improving quality assurance on the shop .

7.3 Data-Driven Yield Improvement

Using MIRAI's analytical workflows, we rapidly pinpointed factors causing yield drops and field failures. The self-serve yield analysis identified patterns in failing units, and the field analysis linked production data to reliability outcomes, providing actionable recommendations. We documented specific cases where addressing a revealed issue (*e.g.*, a particular SOP deviation or a common factor among failing units) led to a measurable increase in first-pass yield. These case studies reinforce the value of moving from reactive problem-solving to proactive, data-driven decision making in manufacturing.

7.4 Failure Prediction Model

We developed and deployed a machine learning model that predicts product failures based on signals including SOP deviations. The model's strong performance in the pilot (catching ~85% of potential failures) underscores the feasibility of predictive quality in manufacturing. Rather than waiting for a failure to occur, the line can now anticipate it and intervene. This represents a shift towards predictive manufacturing operations, where each unit's risk is continually assessed and mitigated in real time.

7.5 Generic and Scalable Framework

Although our implementation was in an electronics assembly context, we designed the framework to be generic. We avoided any reliance on proprietary tools or product-specific heuristics, focusing instead on general techniques (computer vision for action recognition, commonality analysis, supervised learning on process data). This makes our approach applicable to a wide range of manufacturing settings where improving yield and quality is critical. Whether it is circuit board assembly, automotive component production, or any process with manual operations, the core idea remains the same: instrument the process with sensors (vision), collect and analyze the data holistically, and use AI to drive continuous improvement.

7.6 Scientific and Practical Relevance

From an academic perspective, our work bridges the gap between theory and practice by applying state-of-the-art AI (deep learning vision, data mining, ML models) to a practical industrial problem, and demonstrating tangible benefits. We also provided citations to comparable approaches in literature and industry, positioning our contributions in context. For instance, our integrated predictive approach can be seen as a novel extension of both traditional quality control and newer smart factory initiatives.

The successful results of the pilot pave the way for broader deployment. In future work, we plan to scale the system to more stations and more complex assembly scenarios, further validating its robustness. We will also explore advanced modeling techniques to improve prediction and possibly automate corrective responses. Another avenue is to incorporate cost optimization – for example, dynamically deciding whether a flagged unit should be reworked or scrapped based on prediction confidence and economic factors.

The integration of MIRAI, computer vision monitoring, and PROSPECT enforcement represents a powerful strategy for electronics manufacturers seeking to achieve higher yields and near-zero defects. By ensuring that the processes are executed as intended and learning from every deviation, manufacturers can dramatically reduce variability and preempt failures. Our research demonstrates that such an approach is not only technically achievable but highly effective. It embodies a shift from reactive quality control to proactive and preventive quality assurance. The manufacturing line becomes a intelligent system: constantly observing, learning, and improving. This leads to tangible gains in efficiency, product quality, and customer satisfaction. As the manufacturing industry continues to embrace digital transformation, we expect that the methodologies outlined in this paper will inform the next generation of smart factories, where data and AI work hand in hand with human operators to drive excellence in production. Ultimately, the synergy of human expertise, advanced analytics, and real-time vision feedback can unlock new levels of performance and reliability in electronics manufacturing and beyond.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

[1] Falavina M. Maximizing First Pass Yield With AI in Manufacturing. Quality Line, 2025. https://quality-line.com

[2] Vision A. Behavior Analysis Use Case, SOP Compliance Monitoring, ADLINK. ADLINK Technology. https://www.adlinktech.com

[3] Tesfaye K, Silva JV, Nayak HS, et al. Standard Operating Procedure on Yield Gap Decomposition for Use Cases Under Excellence in Agronomy: Understanding Major Yield Drivers for Designing Interventions and Closing Yield Gaps. 2023.

[4] Mody V. Quality in high-volume electronics design: Manufacturing and deployment. Dog Ear Publishing, 2016.

[5] Kalvari N, Lotan N, Zidon M. IT@Intel: Transforming Manufacturing Yield Analysis With AI. White Paper, Intel, 2021. https://www.intel.com

[6] Saihi A, Awad M, Ben-Daya M. Quality 4.0: leveraging Industry 4.0 technologies to improve quality management practices – a systematic review. International Journal of Quality & Reliability Management. 2021, 40(2): 628-650. https://doi.org/10.1108/ijqrm-09-2021-0305

[7] Das S. Computer Vision AI in SOP Monitoring in Manufacturing, 2024. https://intelgic.com

[8] Suma KG, Patil P, Sunitha G, et al. Computer Vision and Its Intelligence in Industry 4.0. Machine Learning Techniques and Industry Applications. Published online May 3, 2024: 119-142. https://doi.org/10.4018/979-8-3693-5271-7.ch007

[9] Duffy JF, Zitting KM, Czeisler CA. The Case for Addressing Operator Fatigue. Reviews of Human Factors and Ergonomics. 2015, 10(1): 29-78. https://doi.org/10.1177/1557234x15573949

[10] IntraStage. How Correlating Failure Analysis and Manufacturing Results Can Help Prevent Future Failures, IntraStage. IntraStage, Apply Manufacturing Intelligence, 2018. https://intrastage.com

[11] Bukhari SMS, Akhtar R. Leveraging Artificial Intelligence To Revolutionize Six Sigma: Enhancing Process Optimization And Predictive Quality Control. Contemporary Journal of Social Science Review. 2024, 2(04): 1932-1948.

[12] Shah D. Leveraging Computer Vision to Tackle Safety and Quality Challenges in Manufacturing. Wevolver, 2025.
https://www.wevolver.com

[13] Gyllenberg J, Nilsson M. Deviation management in high-mix low-volume production: A case study conducted in the defense industry, 2024.

[14] Janecki L, Reh D, Arlinghaus JC. Challenges of Quality Assurance in Early Planning and Ramp Up of Production Facilities - Potentials of Planning Automation via Virtual Engineering. Procedia Computer Science. 2024, 232: 2498-2507.
https://doi.org/10.1016/j.procs.2024.02.068

[15] Doostan M, Chowdhury BH. Power distribution system fault cause analysis by using association rule mining. Electric Power Systems Research. 2017, 152: 140-147.
https://doi.org/10.1016/j.epsr.2017.07.005

[16] Della Corte R. Understanding the Error Behavior of Complex Critical Software Systems through Field Data. Diss. University of Naples Federico II, Italy, 2016.

[17] Romano G, Conti A. The role of Customer Feedback Loops in driving Continuous Innovation and Quality Improvement. National Journal of Quality, Innovation, and Business Excellence. 2024, 1(2): 30-39.

[18] Collins DH, Huzurbazar AV, Warr RL. Highly accelerated life testing (HALT): A review from a statistical perspective. WIREs Computational Statistics. 2024, 16(4).
https://doi.org/10.1002/wics.70000

[19] Chen HH, Hsu R, Yang P, et al. Predicting system-level test and in-field customer failures using data mining. 2013 IEEE International Test Conference (ITC). Published online September 2013.
https://doi.org/10.1109/test.2013.6651892

[20] Tran PH, Ahmadi Nadi A, Nguyen TH, et al. Application of Machine Learning in Statistical Process Control Charts: A Survey and Perspective. Control Charts and Machine Learning for Anomaly Detection in Manufacturing. Published online August 30, 2021: 7-42.
https://doi.org/10.1007/978-3-030-83819-5_2

[21] Cui Y, Kara S, Chan KC. Manufacturing big data ecosystem: A systematic literature review. Robotics and Computer-Integrated Manufacturing. 2020, 62: 101861.
https://doi.org/10.1016/j.rcim.2019.101861

[22] Schönfub B. How AI Is Transforming the Factory Floor. World Economic Forum, 2024.
https://www.weforum.org

[23] Gorelik A. The enterprise big data lake: Delivering the promise of big data and data science. O'Reilly Media, 2019.

[24] Granados Segura L. Enhancing Modeling and Motion Analysis for Industrial Pastry Dough Quality. MS thesis. Universitat Politècnica de Catalunya, 2024.

[25] Camp RC. Benchmarking. Published online October 1, 2024.
https://doi.org/10.4324/9781003578871

[26] Jiang Y, Yin S, Kaynak O. Performance Supervised Plant-Wide Process Monitoring in Industry 4.0: A Roadmap. IEEE Open Journal of the Industrial Electronics Society. 2021, 2: 21-35.
https://doi.org/10.1109/ojies.2020.3046044

[27] Pardo-Calvache CJ, García-Rubio FO, Piattini-Velthuis MG, et al. A 360-degree process improvement approach based on multiple models. Revista Facultad de Ingeniería Universidad de Antioquia. 2015 (77): 95-104.

[28] Qudus L. Leveraging Artificial Intelligence to Enhance Process Control and Improve Efficiency in Manufacturing Industries. International Journal of Computer Applications Technology and Research. 2025, 14(02): 18-38.

[29] Ghelani H. Advanced AI Technologies for Defect Prevention and Yield Optimization in PCB Manufacturing. International Journal of Engineering and Computer Science. 2024, 13(10): 26534-26550.
https://doi.org/10.18535/ijecs/v13i10.4924

[30] Diez-Olivan A, Del Ser J, Galar D, et al. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. Information Fusion. 2019, 50: 92-111.
https://doi.org/10.1016/j.inffus.2018.10.005

[31] Paneerselvam N, Muhammad NA, Azhan AM, et al. Analyzing critical success factors in Lean Six Sigma training. International Journal of Productivity and Performance Management. 2024, 74(4): 1400-1424.
https://doi.org/10.1108/ijppm-11-2023-0627

[32] Wolniak R. The usage of Poka-Yoka in Industry 4.0 conditions. Zeszyty Naukowe. Organizacja i Zarzadzanie/Politechnika Śląska, 2024.

[33] Pinciroli Vago NO, Forbicini F, Fraternali P. Predicting Machine Failures from Multivariate Time Series: An Industrial Case Study. Machines. 2024, 12(6): 357.
https://doi.org/10.3390/machines12060357