Team Details

Team Name: Ghost in the wafer

SR. NO	ROLE	NAME	ACADEMIC YEAR
1	**Team Leader**	*Shireen Mir*	*3rd*
2	Member 1	*Varad Pashte*	*3rd*
3	Member 2	*Bhumika Parulekar*	*3rd*
4	Member 3	*Nigam Mehta*	*2nd*

COLLEGE NAME

Shah And Anchor Kutchhi Engineering College

TEAM LEADER CONTACT NUMBER

+91 90290 92144

TEAM LEADER EMAIL ADDRESS

shireen.17924@sakec.ac.in

Problem Statement Addressed

Semiconductor fabrication involves hundreds of precise steps where even microscopic defects can cause catastrophic yield loss. Modern inspection tools generate terabytes of high-resolution images daily, creating a critical bottleneck: centralized analysis and manual review are too slow, bandwidth-intensive, and prone to error for real-time production needs. To maintain high throughput and secure intellectual property, the industry urgently requires an intelligent Edge-AI solution capable of detecting and classifying defects directly on the equipment with millisecond latency, eliminating the dependency on cloud connectivity or expensive backend infrastructure.

Proposed Solution –

Our Proposed Edge-AI Solution

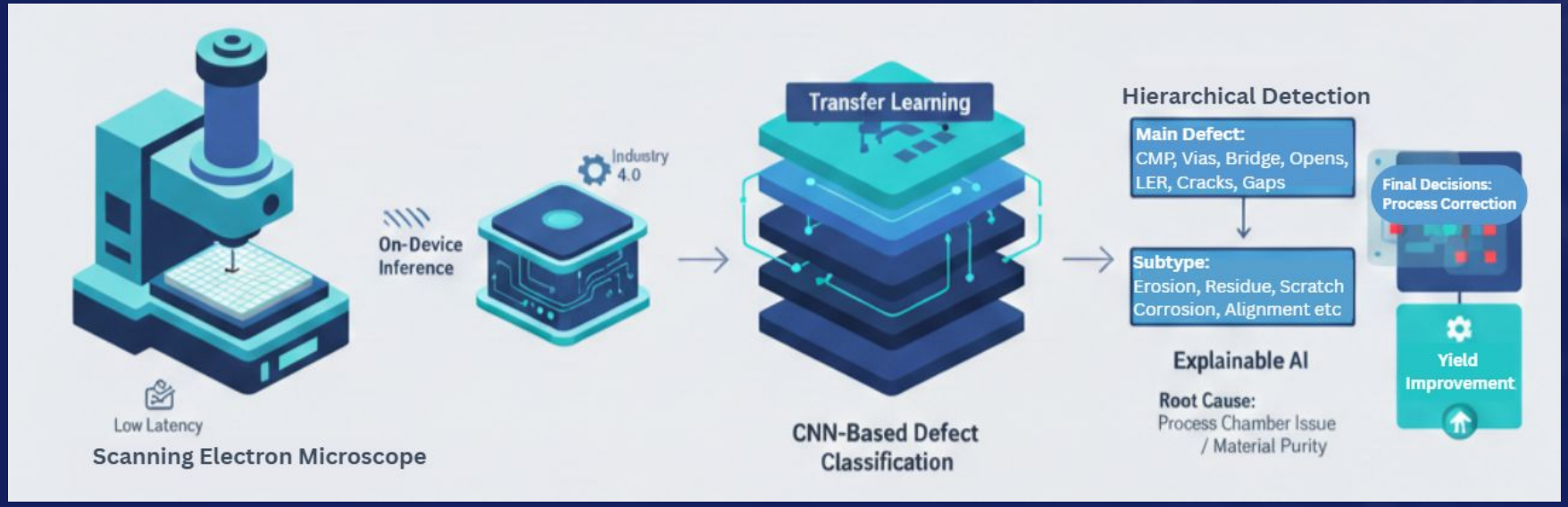
- CNN-based defect classification
- Transfer learning for small datasets
- Hierarchical defect → subtype detection
- Explainable output with root cause

Why Edge-AI?

- Low latency
- On-device inference
- Scalable & Industry-4.0 ready

Inspired From IEEE Research

- CNN for wafer defects
- Transfer learning robustness
- Unknown defect handling
- Hierarchical classification



Innovation and Uniqueness

KEY INNOVATION

Hierarchical AI defect intelligence
→ Classifies defect type first, then identifies exact sub-type

Explainable Edge-AI outputs
→ Generates defect reason, confidence level, and corrective action

Transfer-learning-driven CNN design
→ Achieves high accuracy even with limited semiconductor datasets

Edge-optimized lightweight architecture
→ Designed for real-time inference on low-power devices

Process-aware defect mapping
→ Links visual defects to manufacturing process causes (CMP, VIAS, LER)

Plug-and-play scalability
→ New defect types can be added without retraining entire model

COMPETITIVE ADVANTAGE

Edge-first AI processing
→ Eliminates cloud latency and bandwidth dependency

Hierarchical defect + subtype classification
→ More precise than flat defect classifiers

Explainable AI output
→ Provides defect cause and corrective action, not just labels

Lightweight & deployable model
→ Optimized for low-power edge devices (MobileNet / EfficientNet)

Works with grayscale inspection images
→ Compatible with existing fab imaging systems

Impact and Benefits

Primary Impact

- Enables real-time, on-device defect detection in semiconductor inspection workflows
- Automates identification of defect type and detailed sub-type (CMP, VIAS, LER)
- Provides explainable outputs linking defects to probable process causes
- Reduces dependency on manual review and centralized cloud analysis
- Supports faster corrective actions at the manufacturing line level

Quantifiable Outcomes

- Defect detection latency reduced by 60–80% using edge inference
- Manual inspection and review effort reduced by ~50%
- Overall yield improvement of 2–5% through early defect identification
- Classification accuracy >90% using CNN with transfer learning
- Lightweight models (<20 MB) suitable for edge deployment
- Reduced wafer scrap, rework cost, and cloud infrastructure overhead

Technology & Feasibility/Methodology Used

IMPLEMENTATION STRATEGY

Dataset Plan & Class Design

- Total Images (current/planned): ~1,060 images
- Number of Classes: 8
  - Opens, Bridge, Gap, Cracks, CMP, VIAS,LER, Clean, Other
- Image Type: Grayscale (SEM / Optical microscopy)
- Class Balance: Minimum 50 images per class
- Split: Train 70% / Validation 15% / Test 15%
- Labeling: Manual + Public datasets

Baseline Model (Phase-1)

- Architecture: MobileNet-V2
- Training: Transfer Learning
  (Pretrained on ImageNet, early layers frozen, classifier fine-tuned on semiconductor defect dataset)
- Input Size: 224 × 224
  (Grayscale images replicated to 3 channels)
- Framework: PyTorch
- Model Size: ~13-14 MB
- Accuracy: 85-92 %
- Evaluation: Confusion Matrix
  (Class-wise performance analysis for Clean vs Defect categories)

GitHub & Video Link

GitHub Repository

shireen90290/ghost_in_the_wafer_phase1: Edge-deployable, explainable AI system for real-time semiconductor wafer defect classification using CNNs and transfer learning, enabling low-latency hierarchical defect and sub-type identification on manufacturing lines.

Research and References

References & Citations

Ref 1: Improvement of Multi-Lines Bridge Defect Classification by Hierarchical Architecture in Artificial Intelligence Automatic Defect Classification
URL: Improvement of Multi-Lines Bridge Defect Classification by Hierarchical Architecture in Artificial Intelligence Automatic Defect Classification | IEEE Journals & Magazine | IEEE Xplore

Ref 2: A Convolutional Neural Network for Fault Classification and Diagnosis in Semiconductor Manufacturing Processes Ki Bum Lee, Sejune Cheon, and Chang Ouk Kim
URL: A Convolutional Neural Network for Fault Classification and Diagnosis in Semiconductor Manufacturing Processes | IEEE Journals & Magazine | IEEE Xplore

Ref 3: On Effectiveness of Transfer Learning Approach for Neural Network-Based Virtual Metrology Modeling Seokho Kang
URL: https://ieeexplore.ieee.org/document/8240679