Team Details

Team Name: VisionX

SR. NO	ROLE	NAME	ACADEMIC YEAR
1	**Team Leader**	*DHINEKKA B*	*2023-2027*
2	Member 1	*JAYANTHI VENKATA SREEJA*	*2024-2028*
3	Member 2	*B SHIVANI SUNDARI*	*2024-2028*

COLLEGE NAME

CHENNAI INSTITUTE OF TECHNOLOGY

TEAM LEADER CONTACT NUMBER

+91 8148692063

TEAM LEADER EMAIL ADDRESS

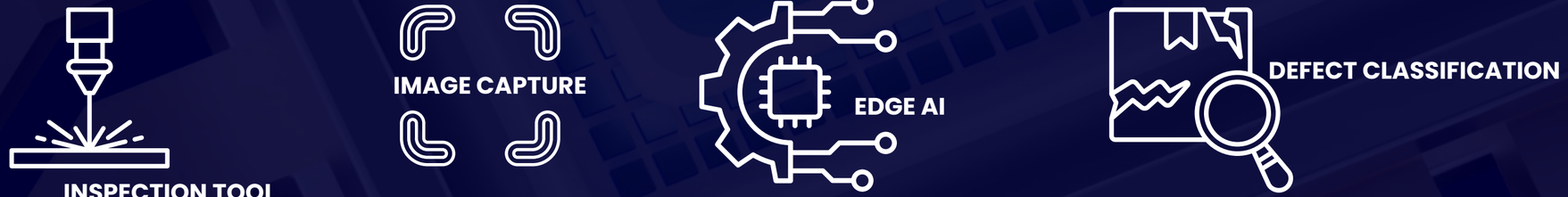
dhinekkab005@gmail.com

PROBLEM STATEMENT ADDRESSED

Edge-AI Based Defect Classification for Semiconductor Wafer/Die Inspection

DESCRIPTION / DETAILS

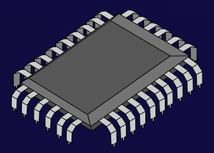
- *Semiconductor fabrication involves hundreds of tightly controlled process steps, where even minor deviations can introduce microscopic defects.*
- *Modern inspection tools (optical/SEM/AFM) generate large volumes of high-resolution wafer and die images, making manual review slow and expensive.*
- *Centralized cloud-based analysis introduces latency, bandwidth bottlenecks, and scalability challenges, especially for high-throughput production lines.*
- *Many defects must be identified in near real-time to prevent yield loss and downstream failures.*
- *There is a strong need for an Edge-AI based defect classification system that can operate close to inspection equipment with low latency and reduced compute overhead.*

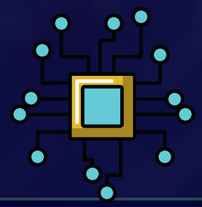
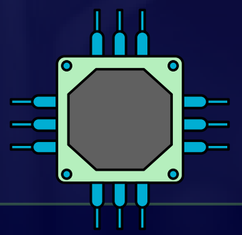


IDEA DESCRIPTION

- *We propose an Edge-AI based image classification system to automatically identify and categorize semiconductor wafer/die defects.*
- *Inspection images captured from fabrication tools are preprocessed and standardized to enable efficient learning under edge constraints.*
- *A lightweight deep learning model is trained to classify images into multiple defect categories, including clean and other cases.*
- *The trained model is designed to be portable to edge platforms, minimizing dependency on centralized cloud infrastructure.*
- *This approach enables low-latency defect detection, supporting real-time decision-making and yield improvement in semiconductor manufacturing.*

Wafer/Die Image → **Preprocessing (Grayscale, Resize)** → **Edge-Friendly CNN Model** → **Defect Class Output**



PROPOSED SOLUTION

- ***Image Acquisition***
  Wafer and die inspection images from optical/SEM inspection systems.
- ***Preprocessing Stage***
  Grayscale conversion and spatial resizing to standardize inputs and reduce computational load.
- ***Defect Classification Model***
  Edge-optimized CNN model trained to classify images into multiple defect categories.
- ***Edge Deployment Readiness***
  Model exported to ONNX format to support deployment on edge AI platforms (e.g., NXP eIQ).
- ***Output & Decision Support***
  Defect class output enables fast inspection feedback and yield monitoring.

Design Principles

- **Edge-first design**
- **Lightweight & portable**
- **Real-time inspection focus**

Why it works

- **Adds conceptual clarity**
- **Reinforces edge-AI theme**
- **Minimal space usage**

TECHNOLOGY & FEASIBILITY/METHODOLOGY USED

IMPLEMENTATION STRATEGY

The proposed solution is implemented using a lightweight deep learning pipeline optimized for edge deployment. The system processes wafer and die inspection images through standardized preprocessing and classifies them using an efficient CNN architecture. The design prioritizes portability, low inference latency, and realistic defect classification suitable for semiconductor manufacturing environments.

Software Architecture

PyTorch-based training and inference pipeline
Folder-based multi-class image classification
ONNX export for edge deployment

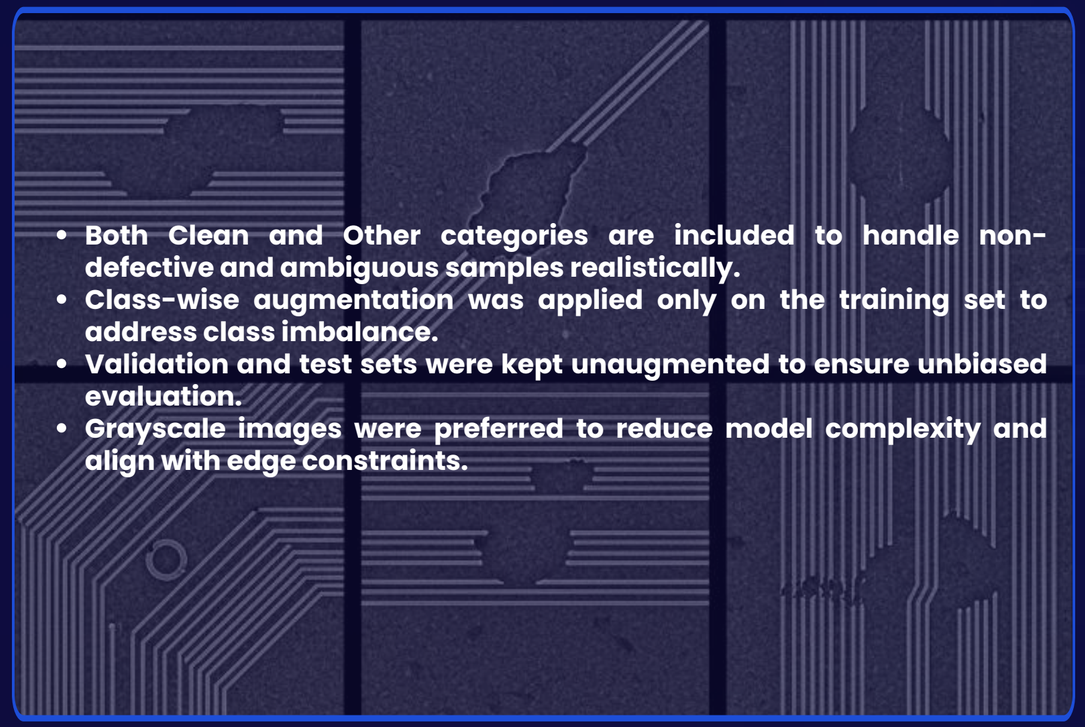
Hardware Components

Targeted for edge-class devices (e.g., NXP i.MX series)
No physical hardware used in Phase-1

Development Tools

Python
PyTorch
OpenCV / PIL
ONNX
GitHub

DATASET PLAN & CLASS DESIGN



- Both Clean and Other categories are included to handle non-defective and ambiguous samples realistically.
- Class-wise augmentation was applied only on the training set to address class imbalance.
- Validation and test sets were kept unaugmented to ensure unbiased evaluation.
- Grayscale images were preferred to reduce model complexity and align with edge constraints.

DATASET SUMMARY

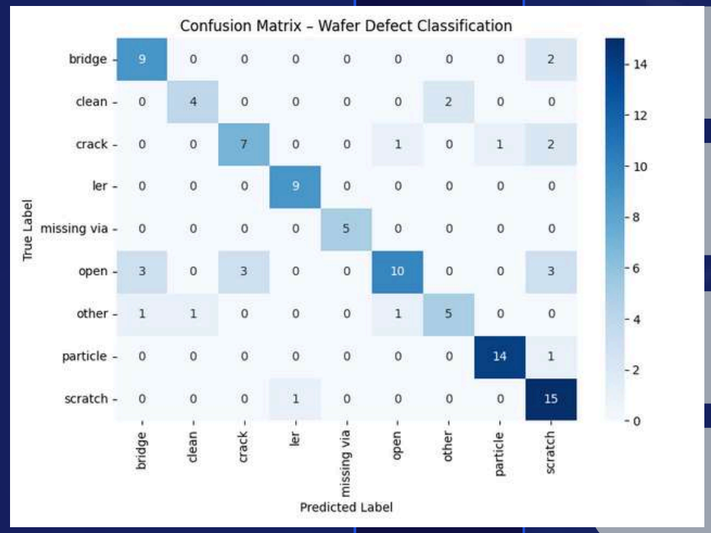
Category	Details
Dataset Size	~1000+ images(with augmentations)
Classes	9 total → 7 defects + Clean + Other
Defect Types	Bridge, Crack, LER, Missing Via, Open, Particle, Scratch
Class Balance	~150 images per class (train set)
Data split	Train / Val / Test = 70 / 15 / 15
Image and Label	Grayscale (160×160), folder-based classification

BASELINE MODEL & RESULTS (PHASE-1)

Test Accuracy: 82%

Macro Precision / Recall / F1: 0.80 / 0.79 / 0.79

9 **No. of Classes: 9**



Architecture: **MobileNetV3-Small**

Training Approach: **Transfer learning**

Input Size: **160 × 160 (grayscale)**
Model size : **307 KB**

Framework: **PyTorch**

Epochs: **20 (best-epoch checkpointing)**

Deployment Format: **ONNX**

ARTIFACTS & LINKS + REFERENCES

GitHub Repository

https://github.com/DhinekkaB/Wafer_Detection.git

DATASET https://drive.google.com/drive/folders/1ELvzk_jxf1cS49J7Bg6TYLVo37f8jUG1?usp=drive_link

Research Background & Methodology

Semiconductor wafer defects manifest as subtle texture and structural variations caused by process deviations during fabrication. Convolutional Neural Networks (CNNs) are well-suited for learning such spatial patterns from inspection images.
To meet real-time manufacturing and deployment constraints, this work adopts an edge-optimized deep learning methodology using grayscale image preprocessing, balanced multi-class defect design, and a lightweight CNN architecture. The approach focuses on robustness, low computational overhead, and unbiased evaluation, making it suitable for practical edge-AI based inspection systems.

References & Citations

Wafer defect inspection using deep learning – survey and industrial case studies

Public SEM / wafer defect image datasets used for academic research

NXP eIQ Edge AI Software Development Documentation