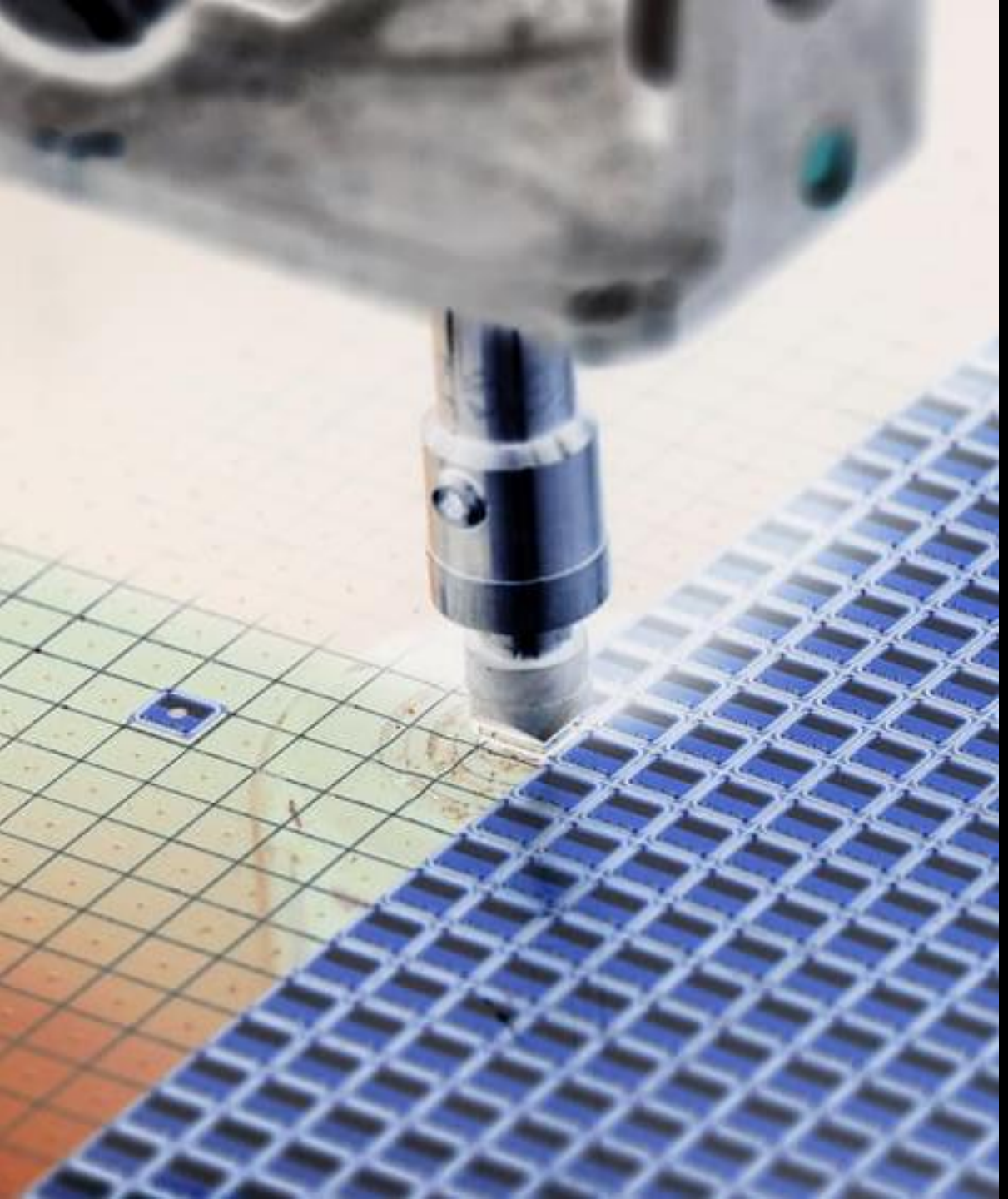
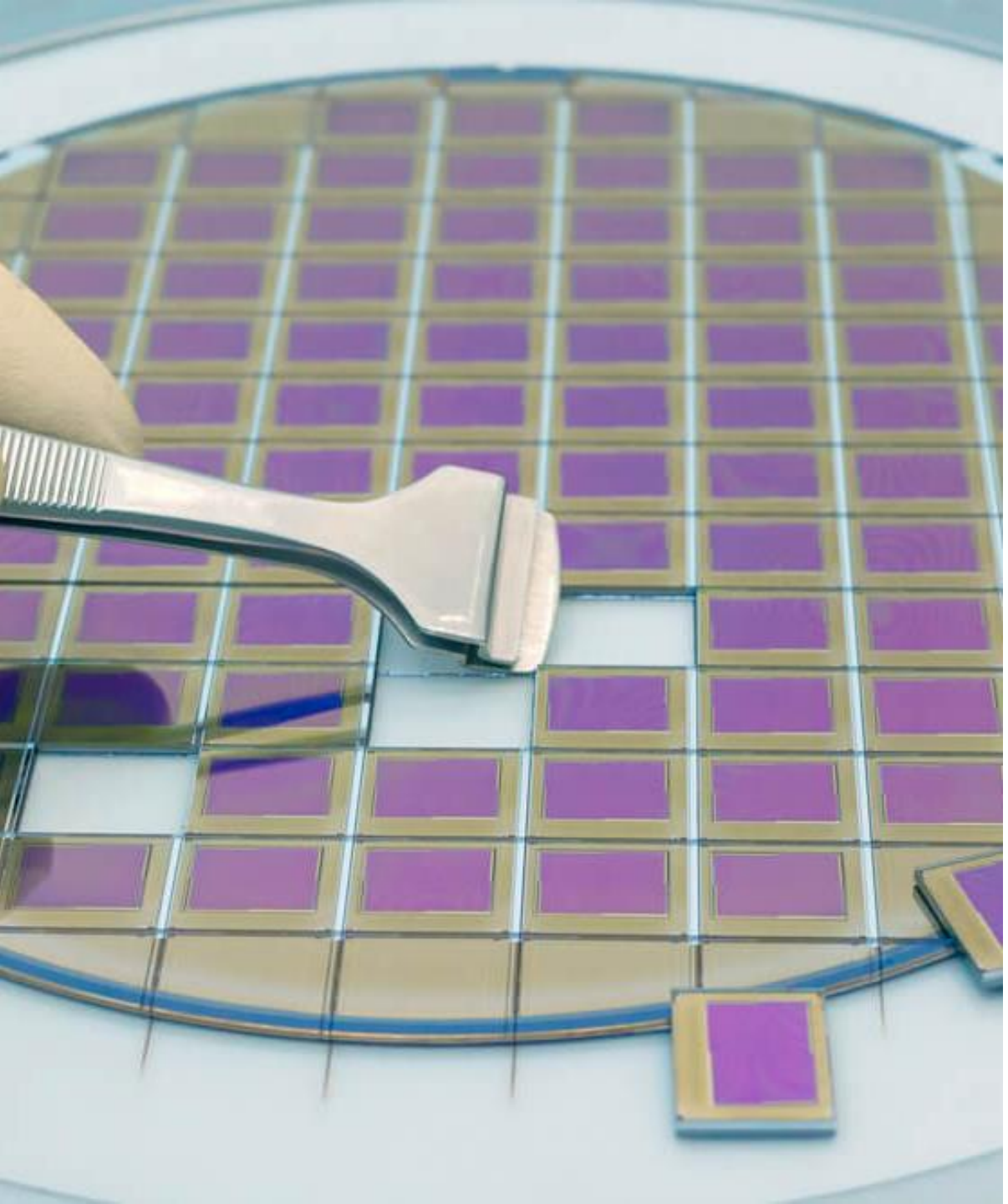
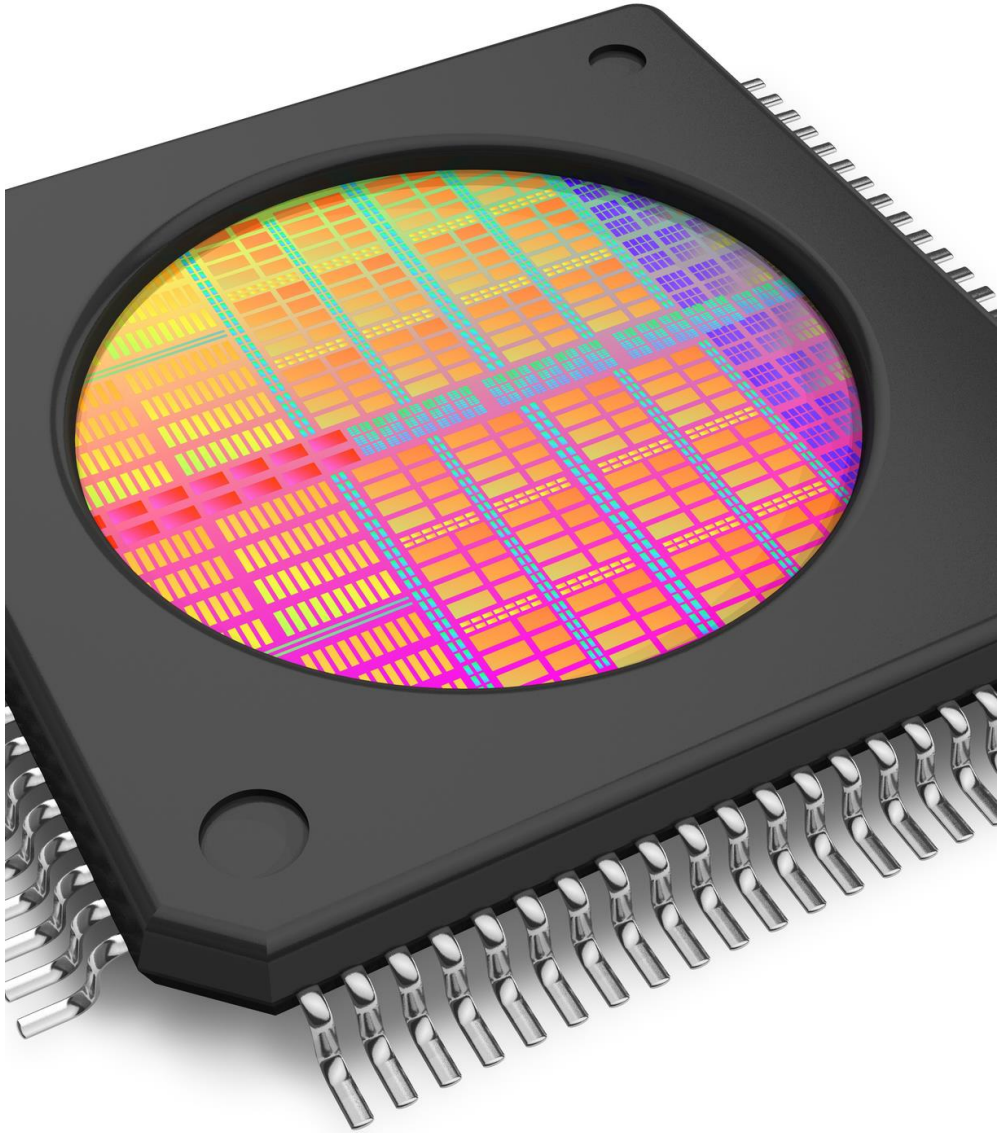


Transforming Semiconductor Yield Management with AI



Group Members

1. Valerie Kigo
2. Lawrence Kamerino
3. Joel Gitonga
4. Hudheyfa Mohamud



CHIP SLEAUTH

An AI-powered wafer defect classification system, a groundbreaking solution that leverages advanced machine learning and computer vision techniques to transform semiconductor yield management.

Introduction & Business Context

Highly complex, capital-intensive semiconductor manufacturing process

Hundreds of fabrication steps involved, with microscopic defects leading to complete product failure and impacting yield and cost

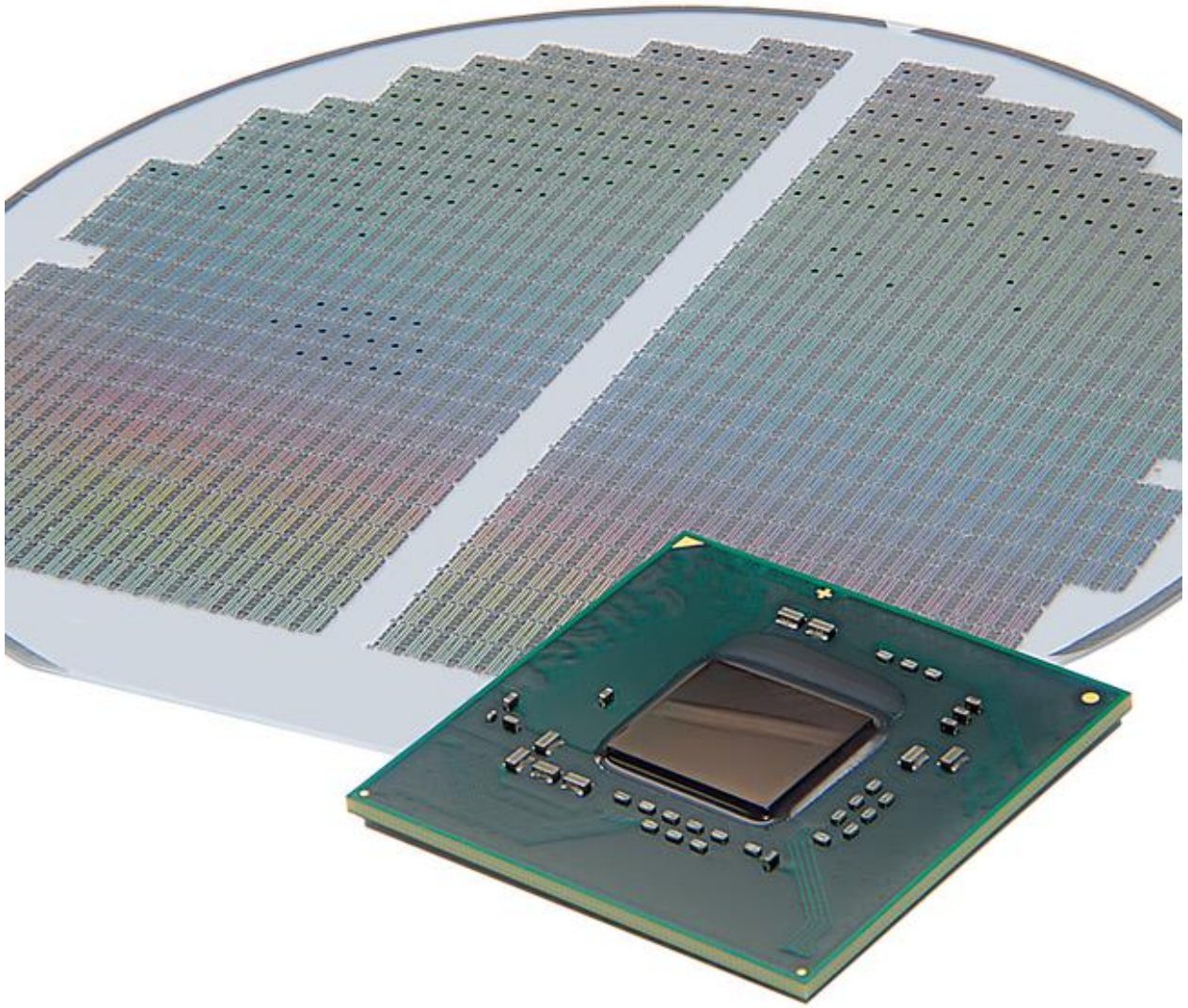
Traditional manual inspection is slow, subjective, and cannot scale with modern production speeds

Inability to keep up with high-volume production and struggles to identify subtle, complex defect patterns

Industry leaders are shifting to AI-driven systems

Leveraging Machine Learning and Computer Vision for early, accurate defect detection to enhance production efficiency

The semiconductor industry is facing challenges with traditional manual inspection methods, presenting an opportunity for AI-driven solutions to transform yield management and production efficiency.



Problem Statement

Manufacturers lack an efficient, automated method to identify and classify wafer defects early in production. Key pain points include manual inspection that fails to scale, inaccuracy and subjectivity in defect identification, and delayed root-cause analysis - leading to reduced yield, increased costs, and longer time-to-market.

Project Objectives & Expected Impact

Primary Objective

Develop and deploy a deep learning-based system to automatically detect and classify wafer defect patterns.

Operational Efficiency

Faster, more accurate defect detection.

Cost Reduction

Reduced labor costs & fewer defective chips.

Quality Improvement

Early detection minimizes yield loss.

Decision Support

Data-driven insights for process optimization.

Scalability

System integrates into production pipelines.

Our Methodology: A Structured Approach

Business Understanding

Analyze the semiconductor manufacturing challenges and identify the opportunities for AI-driven solutions.

Data Acquisition

Gather the necessary data, including the publicly available WM811K wafer defect dataset, to support the project.

Data Preparation & EDA

Clean, standardize, and explore the dataset to gain insights and prepare it for model training.

Feature Engineering

Extract relevant statistical and texture features from the wafer map images to enhance the performance of traditional machine learning models.

Model Training

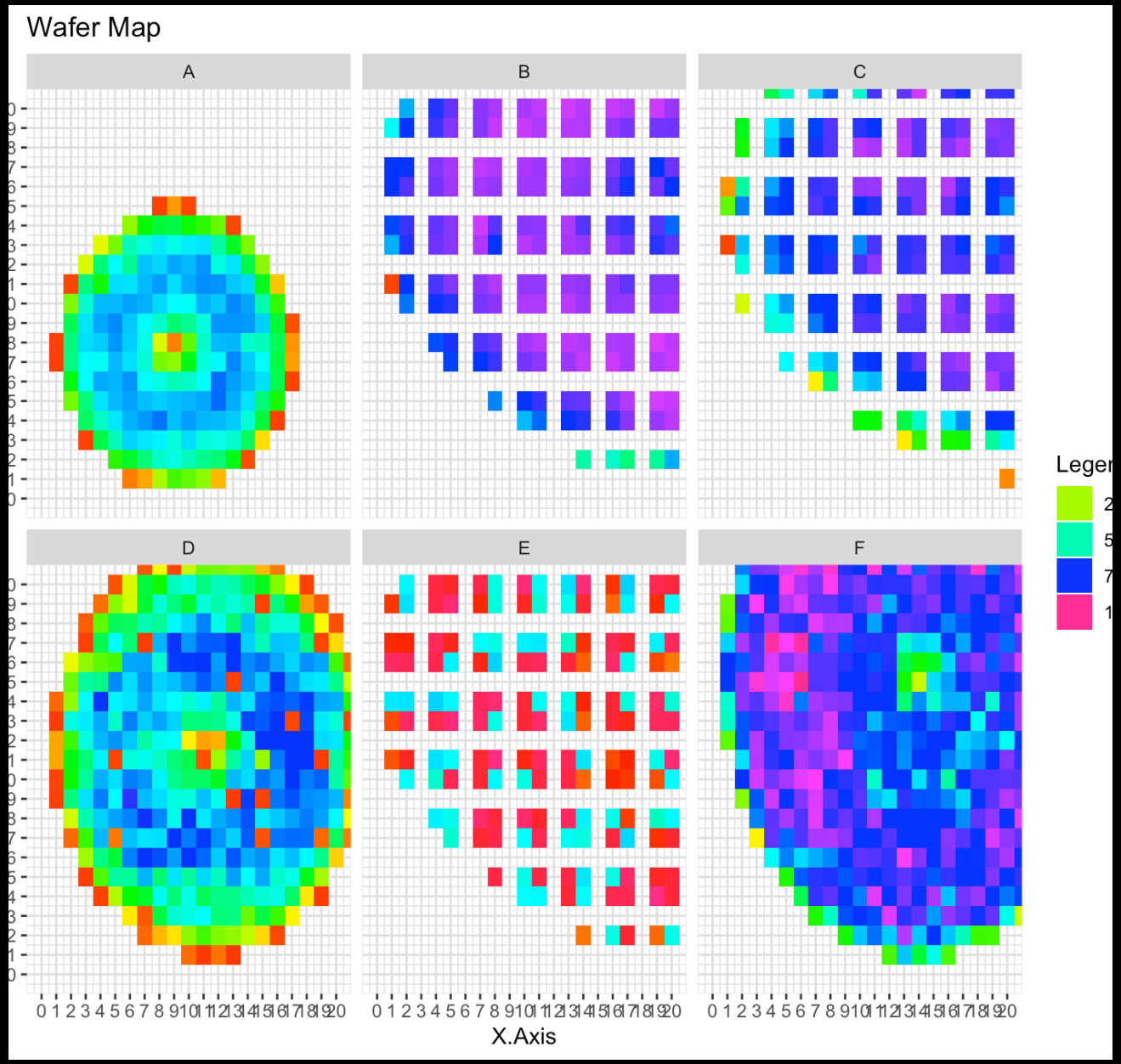
Train multiple machine learning and deep learning models, including Logistic Regression, Random Forest, XGBoost, and Convolutional Neural Network (CNN).

Model Evaluation

Assess the performance of the trained models using appropriate metrics, such as accuracy and F1-score, to identify the optimal solution.

Deployment & Dashboard

Develop a Streamlit-based interactive dashboard to deploy the trained CNN model and provide a tangible tool for process engineers.



Data Understanding: The WM811K Dataset

The WM811K dataset is a publicly available dataset from TSMC that contains over 811,000 wafer samples. Each sample includes a wafer map image and metadata such as defect type and lot name. However, the dataset poses a significant challenge due to severe class imbalance, with over 96% of the samples being 'None' or unlabeled, and the critical defect patterns (Scratch, Donut, Random) making up less than 1% of the data.

Data Preparation & Feature Engineering

Task	Description
Data Cleansing & Standardization	Unified inconsistent labels and filtered out rare classes. Standardized all wafer map dimensions to 32x32 pixels. Ensured data integrity and quality for modeling.
Feature Engineering for Enhanced Models	Extracted statistical and texture features (mean intensity, entropy, edge density). This provided richer input for traditional ML models (Random Forest, XGBoost) beyond raw pixels.
Class Balancing	Applied SMOTE to synthetically generate samples for rare defect types, preventing model bias.

Modeling Approach: A Multi-Model Strategy

Logistic Regression

Established a performance benchmark as the Baseline Model.

Random Forest & XGBoost

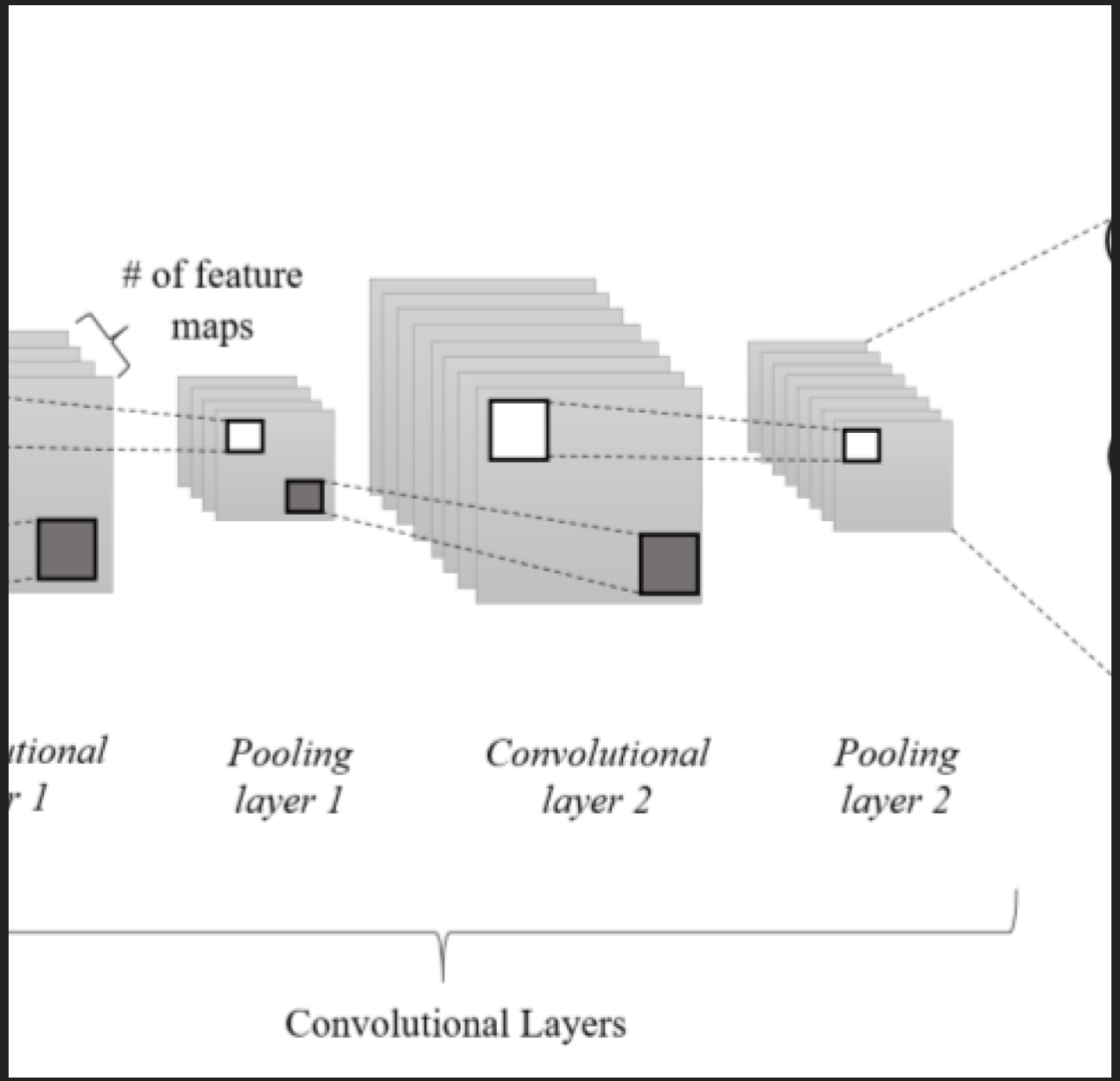
Leveraged engineered features for robust performance as Feature-Based Models.

Convolutional Neural Network (CNN)

Automatically learns spatial patterns from wafer maps as the Image-Based Model.

Model Evaluation & Performance Comparison

Metric	Logistic Regression	Random Forest	XGBoost	CNN
Accuracy	62%	81%	81%	72%



CNN Model Architecture

The CNN model architecture used in this project follows a sequential structure of convolutional, pooling, batch normalization, and dense layers. The convolutional layers are responsible for detecting low-level features such as edges, shapes, and patterns within the wafer map images. The pooling layers then reduce the dimensionality of the feature maps, ensuring spatial invariance. Batch normalization is applied to stabilize the training process and accelerate convergence. Finally, the global average pooling and dense layers transform the learned features into a probability distribution over the defect class labels.

Key Findings & Business Interpretation

AI is Viable

Deep learning can successfully automate wafer defect classification with high accuracy.

CNN is Optimal

For image-based patterns, CNNs outperform traditional ML models, justifying the architectural complexity.

Data Quality is Critical

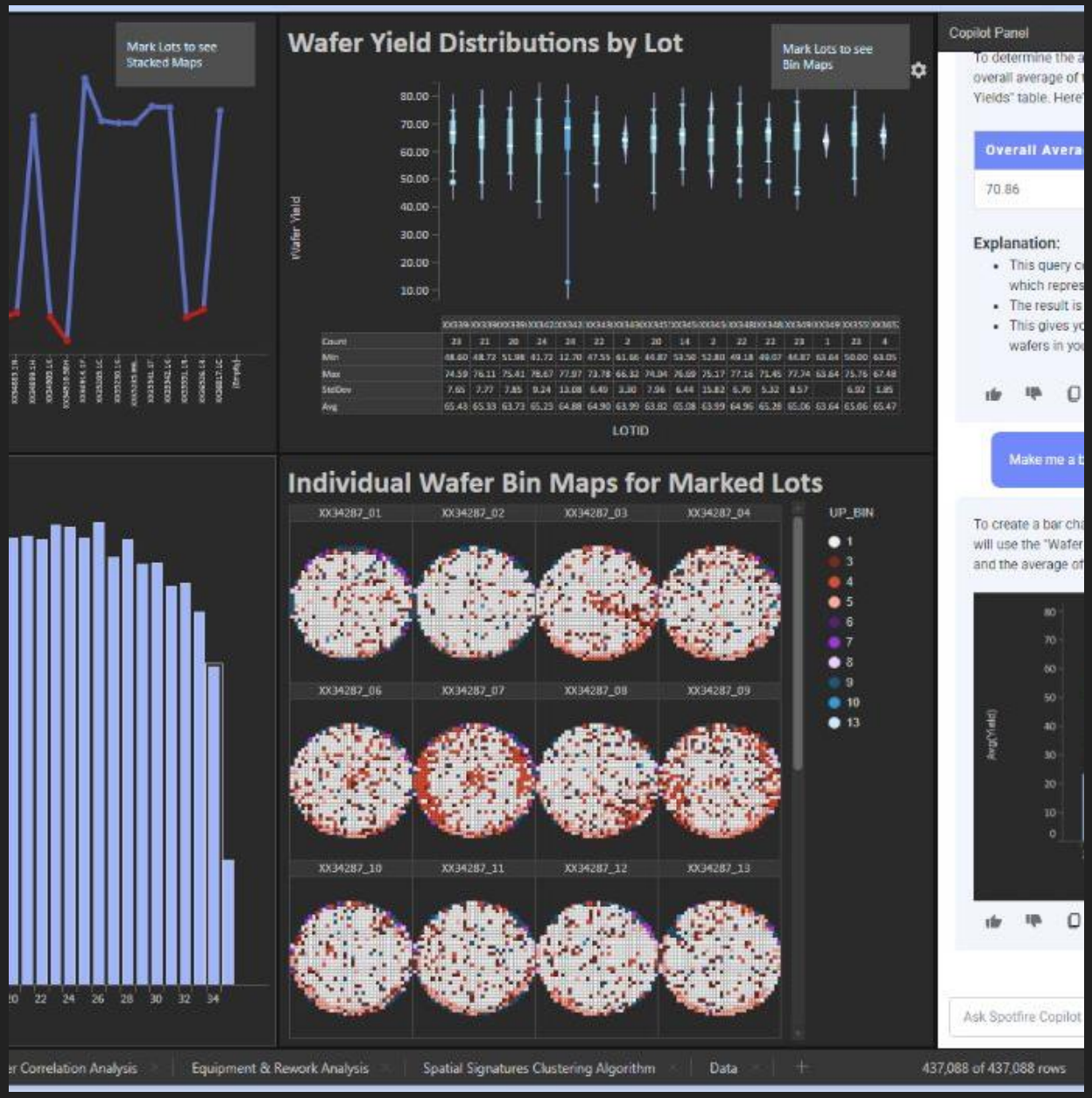
Cleaning, standardization, and balancing were fundamental to our success.

Actionable Insights

The model doesn't just predict; it identifies *where* and *what* the defect is (e.g., Center, Edge-Ring, Scratch), enabling rapid root-cause analysis.

Recommendations

- **Use the CNN model as the main model for wafer defect detection.**
  Leverage the CNN model's superior performance for image-based wafer defect classification.
- **Keep Random Forest or XGBoost as backup models for cases where image data is unavailable or computational resources are limited.**
  Maintain the feature-based traditional models as alternatives when the CNN model cannot be used.
- **Deploy the trained CNN model in the Streamlit app to allow easy image uploads and predictions.**
  Provide a user-friendly interface for process engineers to upload wafer images and get real-time defect classifications.
- **Add Grad-CAM visualizations to explain the CNN's predictions and highlight key defect regions.**
  Enhance the model's interpretability by generating visual explanations for the defect classifications.
- **Continuously update the dataset with new wafer maps to improve model performance over time.**
  Maintain and expand the training data to ensure the model stays up-to-date and adapts to evolving defect patterns.



Deployment: Interactive Defect Classification Dashboard

We have built a functional Streamlit prototype to demonstrate value and gather user feedback. The interactive dashboard allows process engineers to upload wafer map images and receive real-time defect classification predictions from the trained CNN model.

The project successfully developed a robust AI-powered wafer defect classification system that leverages advanced deep learning techniques to address the key challenges in semiconductor manufacturing. By deploying this solution, semiconductor fabs can expect significant improvements in operational efficiency, cost reduction, quality enhancement, and data-driven decision support. The flexible and scalable CNN model demonstrated superior performance, paving the way for widespread adoption of AI in yield management.