Background and Motivation

The semiconductor industry has become increasingly significant, especially with the advent of AI and machine learning technologies. NVIDIA, a leading American semiconductor company, exemplifies this growth, rapidly becoming one of the world’s most valuable companies. As top semiconductor companies like Intel and AMD race to enter the GPU space and capitalize on the AI boom, the demand for high-quality chips and chiplets, processed from wafers through thousands of intricate processes, has surged.

Ensuring pristine cleanliness during wafer processing is crucial, as even minuscule particles can affect the quality of the final product. Despite efforts to maintain a clean environment, controlling particle introduction remains a challenging task. Semiconductor companies develop wafers in controlled environments to minimize airborne particles, particularly within the processing tools. However, periodic maintenance exposes these tools to fab air, potentially introducing contaminants.

Maintaining cleanliness is further complicated by unexpected particle accumulation within the tools. To address this, production tools use dummy silicon wafers, test wafers, to verify air quality. Tools must meet cleanliness standards before processing production-grade wafers. Ideally, test wafers should collect zero particles, indicating a clean air within the tool. In reality, some particle accumulation is inevitable, with a random distribution and a maximum allowable particle count serving as the baseline. Specific particle patterns on test wafers can indicate internal tool contamination, necessitating targeted cleaning. For instance, a robot arm for moving wafers can cause particle accumulation around the wafer’s contact points if the robot arm is dirty. Identifying the particle source based on distribution patterns is vital to minimize equipment downtime as extended tool downtimes reduce wafer processing rates, slowing the production of highly in-demand wafers.

Currently, companies invest in metrology tools sensitive to particle accumulation to aid tool engineers in troubleshooting particle sources, though, machine learning has been introduced to auto-classify potential particle sources to further minimize the troubleshooting time. This project proposes using machine learning for particle classification on semiconductor wafers, beginning with multiclass logistic regression as a baseline model. By exploring hyperparameter optimization and more advanced techniques like CNN, we aim to improve prediction accuracy and correctly classify any noticeable patterns.

Resources

-<https://www.kaggle.com/datasets/qingyi/wm811k-wafer-map>