A Potential Research Proposal Writing with Data Science

A Term Paper

Presented to

MSE 542A Data Science and Materials Informatics

Department of Materials Science and Engineering

University of Washington

In Partial Fulfillment

Of the Requirements of the Degree

Master of Science

By

Owen Chang Chien

Autumn 2022

Section 1 (Description of your selected research topic.)

Crystalline defects play an important role in the field of materials science, since they could drive most materials properties such as optical properties or semiconducting behavior. We can say that to control defects is to control properties. Materials defects could be classified into different types based on their dimension. There are point defects (0-dimensional), linear defects (1-dimensional), planar defects (2-dimensional), and volume defects (3-dimensional). Although defects like cracks or voids could be observed via manual inspection, most of them still need advanced techniques to deal with because of their atomic size. I am interested in the real-time materials defects analysis (the evolvement of defects) and image data processing, and I believe that machine learning could be useful to implement such research.

Section 2 (Why data science or machine learning can add value to your research topic.)

Use Machine Learning to assist the defect analysis may facilitate or even replace manual inspection. With different manufacturing or experimental synthesis methods, we could expect that there are different layouts and types of imperfections in one material. Hence, we can make computer simulations to demonstrate the possible defects inside the material. Then, we could use the data from the simulation to build a machine learning model. For example, we can use Linear Support Vector Machine (SVM) to classify the types of defects. In this way, we could conclude that which manufacturing or experimental synthesis method is the most efficient and effective way under certain condition.

Section 3

In materials science, doping could usually change one material's optical or electrical properties depending on dopant atoms' number or categories. For example, gemstones like sapphire could display several different colors with different dopants such as magnesium or titanium. As for my research proposal, I plan to dope different elements or compounds into semiconductors materials or ceramic materials to see the progression of defects and their final layout. By controlling the defects layout, we could possibly modify the materials' properties and get a better material performance.

In a nutshell, the research could be divided into four steps. First, I would like to use diffusivity dataset such as the one in the references section [1] to build a machine learning model for prediction and classification. I want to prove that the prediction is compatible with the result from imaging techniques such as TEM, STM, or AFM. Then I would implement the doping process and use the imaging techniques to get the images of defects. After getting enough images as my dataset from the imaging techniques, I plan to use Convolutional Neural Networks (CNN) to process the image data and to locate the positions of the defects. Through image processing, we could increase the accuracy of defect characterization. Finally, we check if the prediction is accurate or not. If the prediction is not as good as it should be, we would need to rebuild the machine learning model to acquire the most accurate one.

Section 4

As stated in section 3, the research proposal could be divided into 4 steps, and these steps involve doping, machine learning prediction, using image techniques, and image processing. The proposal could help pave a path for my future career. For instance, in semiconductor industry, an analysis of the wafer map is helpful for quality control and process evaluation in the semiconductor manufacturing process. The defects on the map are mainly caused during the semiconductor device

fabrication process (fab process) [2]. There are two main types of defects on a wafer, random defects and systematic defects [3]. The defects could be predicted and detected much more precisely using the machine learning model. Using such tools could help engineers find abnormal faults more quickly in the production line, improve the efficiency of integrated circuit production, and reduce the rejection rate [4].

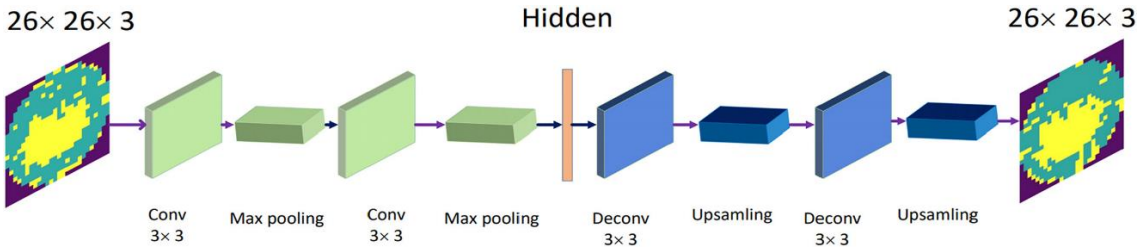

Figure 1 Network structure of convolutional autoencoders for wafer data augmentation [4]

References

1. https://www.ctcms.nist.gov/~gkl/overview.html

2. Kim, E., Choi, S., Lee, D., Kim, K., Bae, Y., Oh, Y. (2021). An oversampling method for wafer map defect pattern classification considering small and imbalanced data. Computers & Industrial Engineering, 16. https://doi.org/https://doi.org/10.1016/j.cie.2021.107767

3. Semiconductor Manufacturing (https://www.hitachi-hightech.com/global/en/knowledge/semiconductor/room/manufacturing/)

4. Yu, N., et al.: Wafer map defect patterns classification based on a lightweight network and data augmentation. CAAI Trans. Intell. Technol. 1– 14 (2022). https://doi.org/10.1049/cit2.12126