Semiconductor Wafer Particle Defect Classification Based on Deep Learning and Multimodal Feature Fusion

Wenxin Liu，Chengjie Bai*
School of Physics and Electronic Sciences,Shandong Normal University,Jinan,Shandong
*E-mail：sdnuliuwenxiny@163.com

1 Abstract

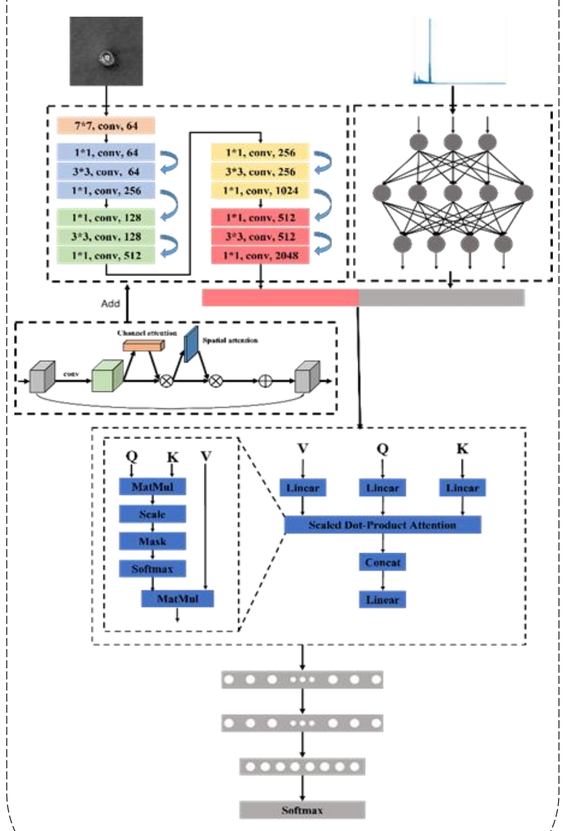
In the semiconductor manufacturing process, the identification and classification of the chemical composition of particle defects on wafers is critical to ensuring product quality. In this study, we propose a deep convolutional neural network (CNN)-based automatic defect classification model that fuses scanning electron microscope (SEM) images and energy dispersive X-ray (EDX) spectral data for more efficient and accurate defect identification.Our model first feeds the image data into a network with Residual Neural Network (ResNet-50) as the model backbone structure and introduces a Convolutional Block Attention Module (CBAM) attention mechanism to the network, uses a multilayer perceptron (MLP) for feature extraction of energy dispersive X-ray (EDX) spectra, achieves feature fusion of image features and spectral features through a multi-head self-attention mechanism, and finally achieves eight classifications of different types of defects in semiconductor particles by means of a fully-connected layer.

2 Introduction

Semiconductor equipment industry has maintained a rapid development trend driven by the domestic market. In the past few years, deep learning technology has made significant progress, and the classification of semiconductor defects has thus gained new advantages, opening up new possibilities for eliminating the source of defects. Classification based on surface defects etc. is currently being investigated by many. Wafers with unknown defects will lead to high raw material costs and high manufacturing precision, ultimately leading to a decrease in yield, and therefore, detection of unknown patterns or anomaly detection is another important aspect. However, there is a paucity of literature on the classification of defect chemical composition.We propose an attention mechanism based defect recognition model for semiconductor particles during semiconductor fabrication that accurately captures and classifies various defects.

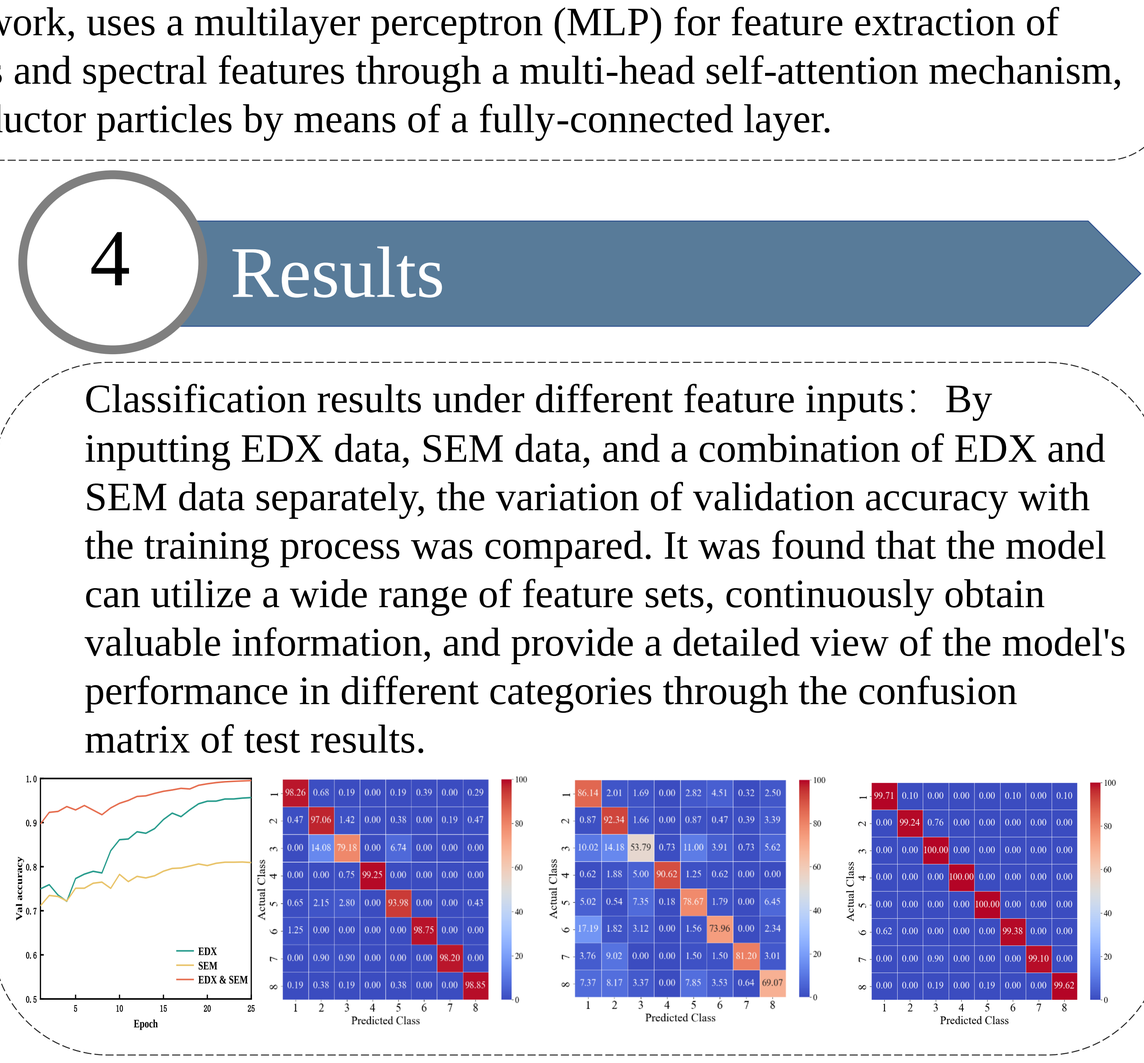
3 Methods

We propose an attention mechanism-based semiconductor particle defect recognition model for semiconductor manufacturing process. The inputs are semiconductor wafers containing particle defects in SEM images and EDX spectra, respectively. A CBAM attention mechanism is introduced to Resnet50 to enhance the network's ability of sensing and learning the key features, by adding Bayesian optimisation hyperparameter optimisation, the network can adjust the model configuration more intelligently, but since the distribution of the dataset is highly heterogeneous, we pre-processed the data to be able to accurately capture and classify various defects.

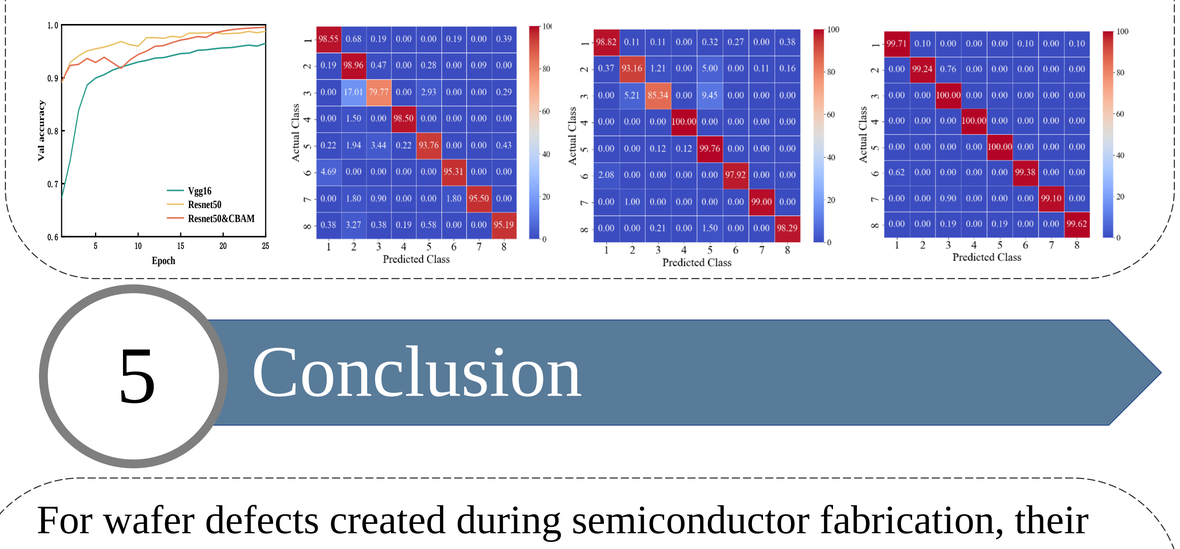


4 Results

Classification results under different feature inputs： By inputting EDX data, SEM data, and a combination of EDX and SEM data separately, the variation of validation accuracy with the training process was compared. It was found that the model can utilize a wide range of feature sets, continuously obtain valuable information, and provide a detailed view of the model's performance in different categories through the confusion matrix of test results.



Classification results of different models: We compared three models: the classic Vgg-16, Resnet50, and the enhanced Resnet50&CBM model. The evaluation of model performance is achieved through detailed analysis of accuracy and confusion matrix on the validation set.



5 Conclusion

For wafer defects created during semiconductor fabrication, their chemical composition can be classified using the model proposed in this paper. The model combines images from a scanning electron microscope (SEM), and spectral data created by energy dispersive X-rays (EDX). By introducing the attention mechanism in Resnet50, extracting the features by using MLP for EDX, and finally combining SEM and EDX to achieve feature fusion for classification through the multi-head self-attention mechanism, the model also overcomes the limitation of uneven data distribution, and Bayesian hyper-parameter optimization has been applied to the proposed model, which improves the efficiency of the model tuning and the classification speed. And it is shown that the average classification accuracy of the model reaches more than 98%.