MATH 6380P Project 2: Nexperia Semi-conductor Classification

Jimmy WU, Zhaoyu SU

Dept. of Civil & Environmental Engineering, HKUST

Introduction & Objectives

In this project, a simple DCNN architecture is applied for semi-conductor defect detection, and the defect location is localized via Gradient-weighted Class Activation Mapping (Grad-CAM) for better visualization and assessment.

Dataset

The dataset used in this study is provided by Nexperia. Some key features of the dataset is:

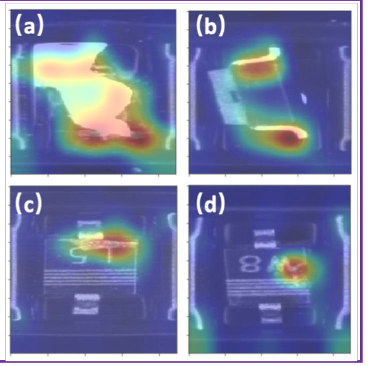
- 224 x 224 greyscale image.
- Highly imbalanced dataset: 27,000 images for good quality while only 3,000 for defect.
- Some defects are hard for detection, even if for human.

Visualization & Conclusion

With the help of the Grad-CAM[1], the attention heatmap w.r.t. the defects can be visualized based on the raw input image.

In the right image, there are 4 defects with representativeness. (a) and (b) are obvious defects, and the heat map matches with the intuition quite well. (c) and (d) correspond to two implicit defects, basically the scratches, well the DCNN is still able to capture them.

The final accuracy we achieved based on the Kaggle testing dataset is 95.8% when the weighted loss strategy were applied, and 94.6% with the dataset resampling strategy.



Special Issues

Preprocessing

Followed the rule of thumb, all the input images are **normalized** by subtracting its mean and dividing by its standard deviation batch wised.

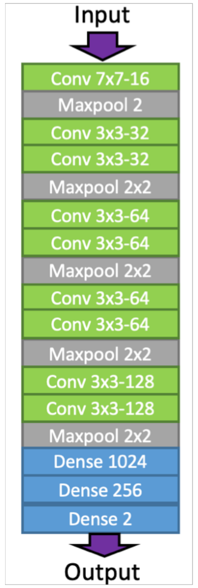
Augmentation are also applied to the input image, including vertical and horizontal flipping, rotating, shearing and shifting.

Imbalance Issue

As the dataset is inherently highly unbalanced, two strategies are adopted to alleviate this issue:

- Resample the dataset with 3,000 defects and 3,000 good quality semi-conductor.
- Add weight to the categorical loss function. (defect: good = 9: 1)

Model Architecture



The DCNN model applied in this project is quite simple (as shown in the left), since the classification has only two classes.

The model is composed of several convolution blocks, while there are two convolution layer within each block, followed by a maxpooling layer.

The popular ResNet-50 [2] is also tested in this project, however, there are no obvious differences observed. As a result, from the perspective of computation power, the ResNet-50 is not chosen at last.

We also observed that the initial learning rate setting in Keras is no longer valid for this application, maybe that's because of the similarity between the two classes. So, instead of the initial learning rate 0.001, we use 0.0001 as our initial learning rate with a decay of 1e-6.

Reference :

1. Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra. "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization." arXiv preprint arXiv:1610.02391 (2017).
2. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).