Wafer Map Failure Pattern Classification Using Deep Learning

Video Link: https://youtu.be/AW3U1vzpIoM

Jie Gong (jieg@stanford.edu), Chen Lin (chenlin1@stanford.edu)

Introduction

Wafer inspection is very important for increasing the yield of a micro/nano-fabrication process in the semiconductor industry because it is possible to figure out the root causes of various process issues based on different kinds of detected wafer map failure patterns [1-4]. The traditional visual recognition approach performed by an experienced person can be expensive and time-consuming [1-4]. In this work, novel deep learning methods are proposed to automate accurate identification of various defect patterns on wafers. The input to our models is a normalized 1-channel wafer map image ($42\times42\times1$) with only one failure pattern from the 8 defect types, and we then used both simplified AlexNet and simplified VGG16 models to output the predicted defect pattern of this wafer map.

Data and Features

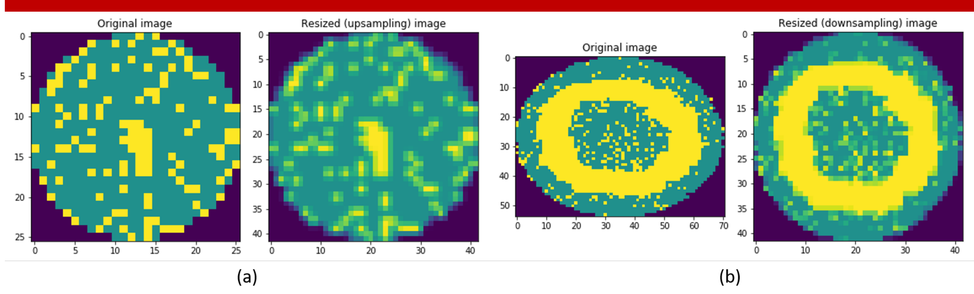


Figure 1. Wafer map comparisons before and after image (a) upsizing or (b) downsizing.

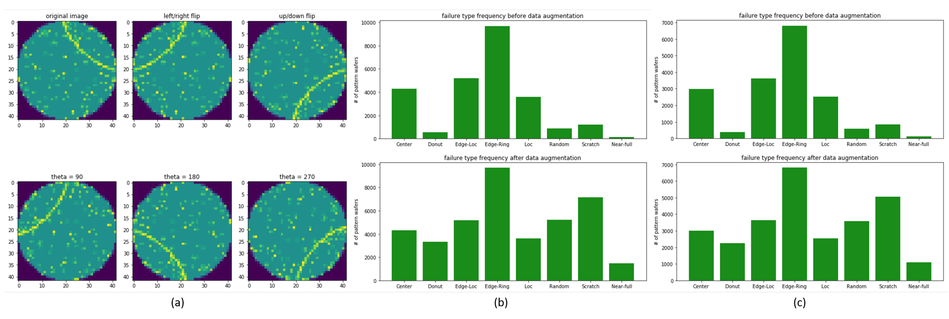


Figure 2. (a) Wafer map comparisons as well as failure pattern type distributions using (b) Approach 1 and (c) Approach 2 before and after data augmentation.

- ❑ Data information
  - ➢ data source: public WM-81K(LSWMD) from Kaggle [5]
  - ➢ Total useful sample size: 25519
  - ➢ 8 classes of failure patterns
- ❑ Data normalization
  - ➢ $42\times42\times1$
- ❑ Data augmentation
  - ➢ flipping and rotating
- ❑ Data split
  - ➢ training : testing = 7: 3
  - ➢ Approach 1: data augmentation after data split
  - ➢ Approach 2: data augmentation before data split

Models

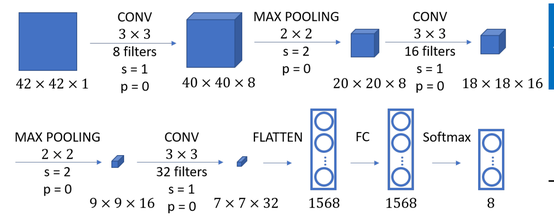


Figure 3. Simplified AlexNet [6] model architecture.

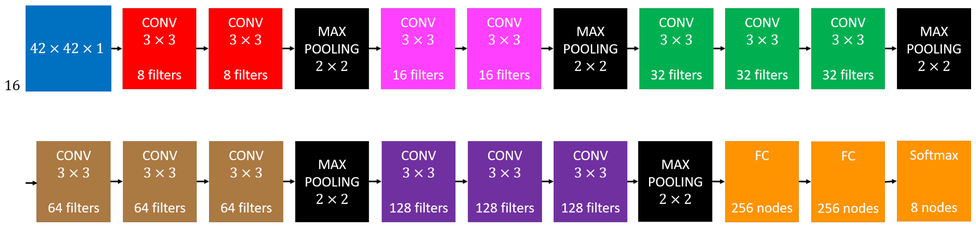


Figure 4. Simplified VGG16 [7] model architecture.

Results and Discussion

- ❑ L2 and dropout regularizations
- ❑ Hyperparameter tuning

Table 1. Bias and variance metrics of all the models.

Model	Data Split Approach and Regularization	Training Accuracy	Testing Accuracy	Variance
Simplified AlexNet	Approach 1 (no regularization)	96.5%	88.5%	8.1%
	Approach 1 (L2 regularization with lambda = 0.001)	93.8%	88.9%	4.9%
	Approach 1 with L2 and dropout regularizations (lambda = 0.001 and keep rate = 0.5)	92.6%	90.6%	2.0%
	Approach 2 (no regularization)	98.5%	88.2%	10.4%
	Approach 2 with L2 regularization	96.5%	89.2%	7.3%
	Approach 2 with L2 regularization plus dropout	92.0%	89.7%	2.3%
Simplified VGG16	Approach 1 (no regularization)	99.8%	92.0%	7.8%
	Approach 1 with dropout regularization (keep rate = 0.5)	98.3%	91.5%	6.8%
	Approach 2 (no regularization)	Similar to random guessing accuracy due to no convergence to a global minimal		
	Approach 2 with dropout regularization (keep rate = 0.5)			

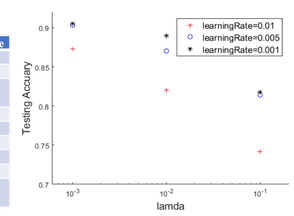
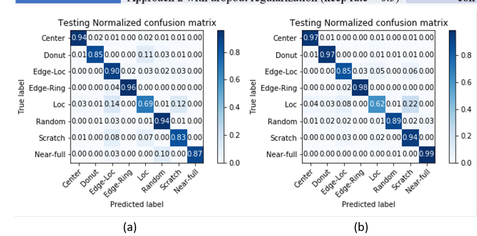


Figure 6. Testing accuracy results for hyperparameter tuning (keep rate = 0.5).



Defect Type	Recall	Precision	F1 Score
Center	96.60%	94.30%	95.40%
Donut	96.50%	92.20%	94.30%
Edge-Local	85.50%	85.10%	85.30%
Edge-Ring	97.70%	97.90%	97.80%
Local	62.00%	79.50%	69.60%
Random	89.50%	97.40%	93.30%
Scatch	94.10%	85.10%	89.30%
Near-full	98.90%	91.00%	94.80%

(c)

Figure 5. Testing confusion matrices of (a) the simplified AlexNet and (b) VGG16 models with regularization. (c) Performance metrics for our simplified VGG16 model with no regularization.

Future Work

- ❑ To solve the non-convergence problem of the simplified VGG16 model using Approach 2
- ❑ To use transfer learning to classify wafer map defect types

References

[1] Wu, Ming-Ju, Jyh-Shing R. Jang, and Jui-Long Chen, "Wafer map failure pattern recognition and similarity ranking for large-scale data sets." *IEEE Transactions on Semiconductor Manufacturing* 28.1 (2014): 1-12.
[2] Fan, Mengying, Qin Wang, and Ben van der Waal. "Wafer defect patterns recognition based on OPTICS and multi-label classification." *2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*. IEEE, 2016.
[3] Piao, Minghao, et al. "Decision Tree Ensemble-Based Wafer Map Failure Pattern Recognition Based on Radon Transform-Based Features." *IEEE Transactions on Semiconductor Manufacturing* 31.2 (2018): 250-257.
[4] Yu, Jianbo, and Xiaolei Lu. "Wafer map defect detection and recognition using joint local and nonlocal linear discriminant analysis." *IEEE Transactions on Semiconductor Manufacturing* 29.1 (2015): 33-43.
[5] https://www.kaggle.com/qingyi/wm811k-wafer-map
[6] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.
[7] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).