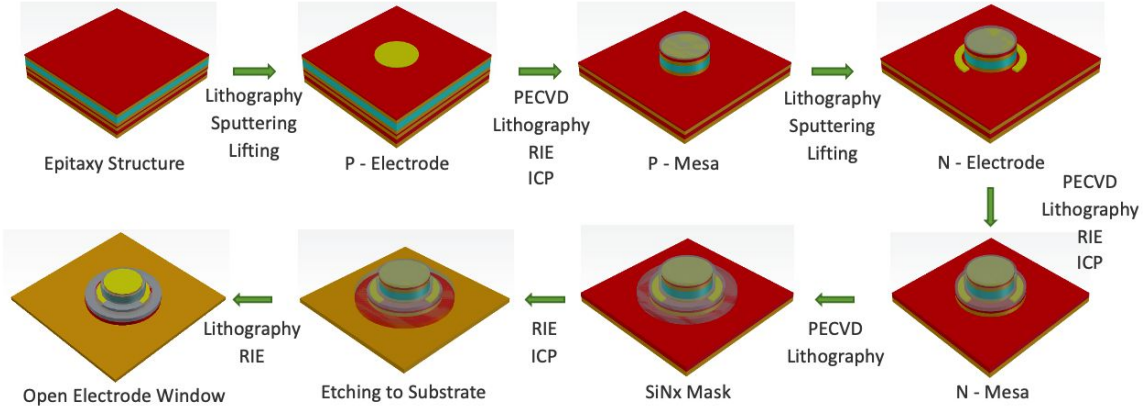
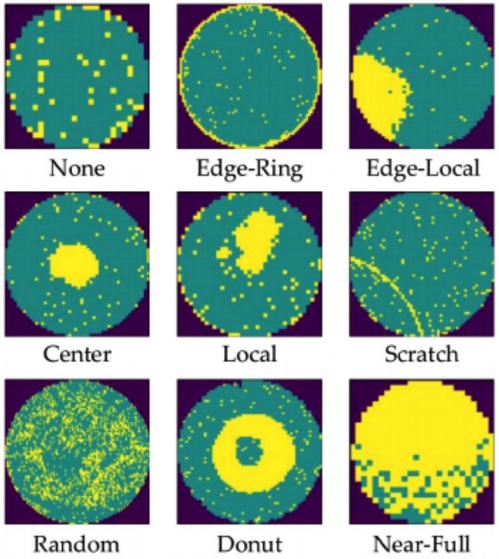
WM-811k Wafermap Defect Classification

Group Member:
Jih-Chin Chen
Ruifeng Luo

Motivation

Project Goal: To identify different types of wafer map defect pattern automatically using cutting-edged deep learning technique instead of manual work, so that the semiconductor fabrication can improve on the yield of wafers manufacturing process.

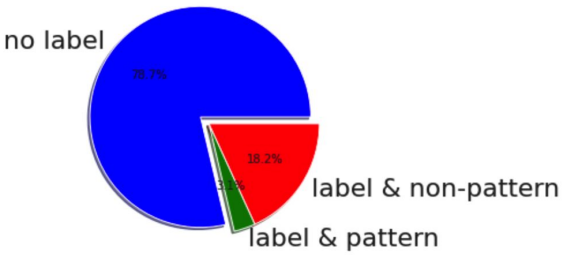


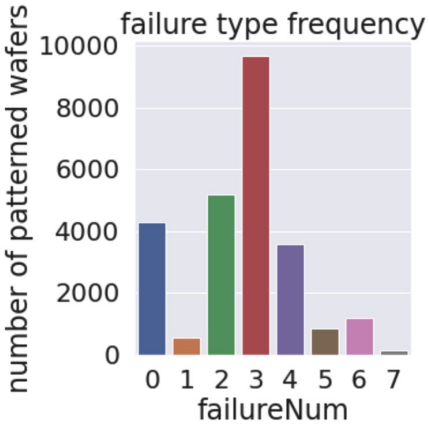


EDA (Exploratory Data Analysis)

Data & Label Distribution

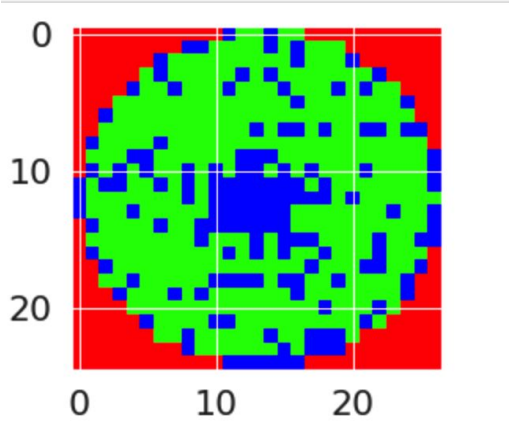
Data with label & pattern only takes only 3.1% of total data and the distribution of label for different failure is imbalance





Resize Input Image with 3 channel

Resize different size image with same dimension and use (R, G, B) channel represent different area.

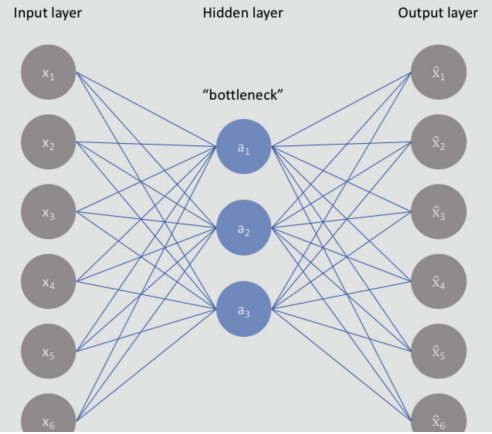


Data Augmentation

AutoEncoder to Oversampling

Train a AutoEncoder model to oversampling minority class adding Gaussian noise.

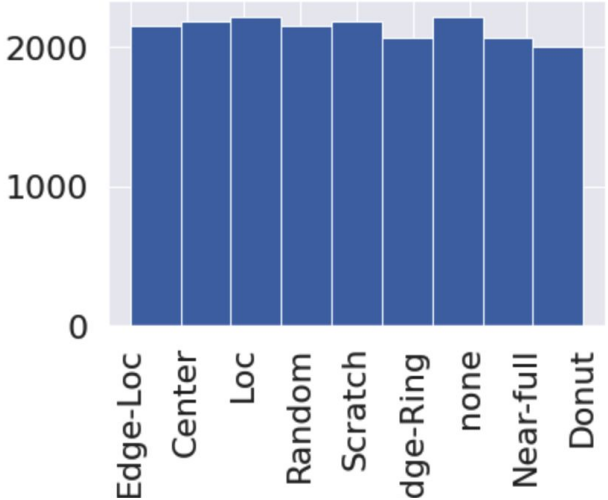
Input / Output: Images
Loss Function: MSE




```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1            [16, 16, 54, 54]             448
              ReLU-2            [16, 16, 54, 54]               0
         MaxPool2d-3            [16, 16, 27, 27]               0
            Conv2d-4            [16, 32, 27, 27]           4,640
              ReLU-5            [16, 32, 27, 27]               0
         MaxPool2d-6            [16, 32, 13, 13]               0
   ConvTranspose2d-7            [16, 16, 26, 26]           2,064
              ReLU-8            [16, 16, 26, 26]               0
   ConvTranspose2d-9             [16, 3, 52, 52]             195
          Sigmoid-10             [16, 3, 52, 52]               0
================================================================
Total params: 7,347
Trainable params: 7,347
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.53
Forward/backward pass size (MB): 23.79
Params size (MB): 0.03
Estimated Total Size (MB): 24.35
----------------------------------------------------------------
```

Random Undersampling

Random select majority class to undersampling.



Convolution Network Model

1. **Feature Learning Layers**

   Convolution + RELU + MaxPool

2. **Classifier (Fully Connected Layers)**

   Linear Layers + RELU + SoftMax

3. **Loss Function**

   Cross Entropy Loss

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1           [16, 32, 56, 56]             896
              ReLU-2           [16, 32, 56, 56]               0
         MaxPool2d-3           [16, 32, 28, 28]               0
            Conv2d-4           [16, 64, 28, 28]          18,496
              ReLU-5           [16, 64, 28, 28]               0
         MaxPool2d-6           [16, 64, 14, 14]               0
            Conv2d-7          [16, 128, 14, 14]          73,856
              ReLU-8          [16, 128, 14, 14]               0
         MaxPool2d-9            [16, 128, 8, 8]               0
           Linear-10                 [16, 1250]      10,241,250
           Linear-11                 [16, 1250]      10,241,250
             ReLU-12                 [16, 1250]               0
           Linear-13                    [16, 9]          11,259
================================================================
Total params: 20,587,007
Trainable params: 20,587,007
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.57
Forward/backward pass size (MB): 48.93
Params size (MB): 78.53
Estimated Total Size (MB): 128.04
----------------------------------------------------------------
```

Result

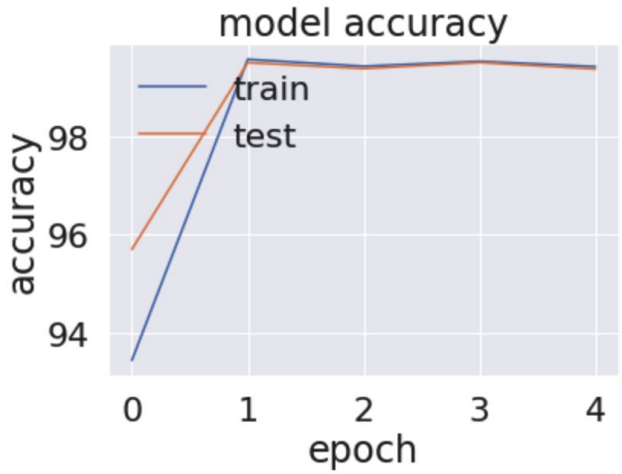
What went well ?

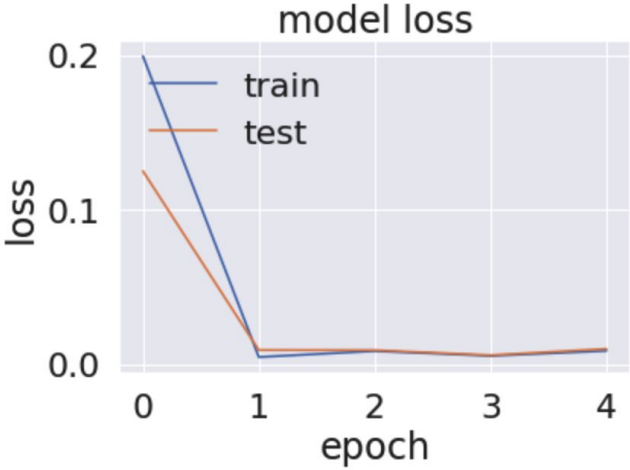
Accuracy: 99.61%

Test Loss: 0.002 (CEL)

What didn’t went well ?

But, cost 6 hours for 10 Kfold training process





Relevant Paper

[1] Tsai, Tsung-Han, and Yu-Chen Lee. "A light-weight neural network for wafer map classification based on data augmentation." IEEE Transactions on Semiconductor Manufacturing 33.4 (2020): 663-672.

[2] Hu, Hanbin, Chen He, and Peng Li. "Semi-supervised Wafer Map Pattern Recognition using Domain-Specific Data Augmentation and Contrastive Learning." 2021 IEEE International Test Conference (ITC). IEEE, 2021.

[3] C. -Y. Wang and T. -H. Tsai, "Defect Detection on Wafer Map Using Efficient Convolutional Neural Network," 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), 2021, pp. 1-2, doi: 10.1109/ICCE-TW52618.2021.9603145.