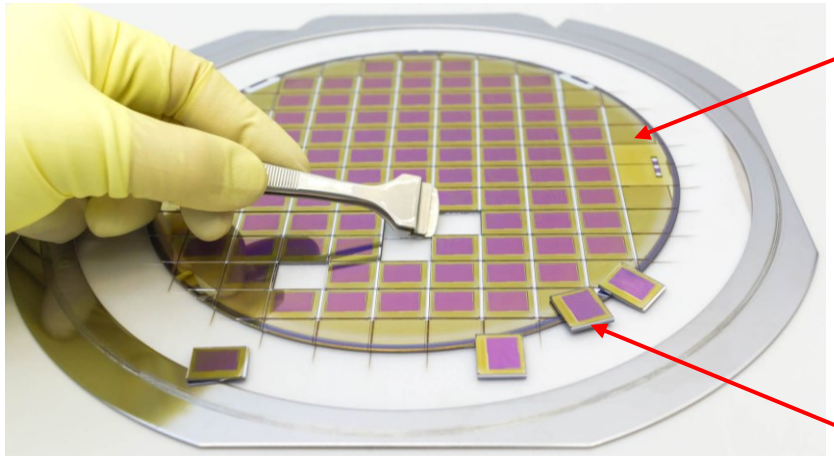
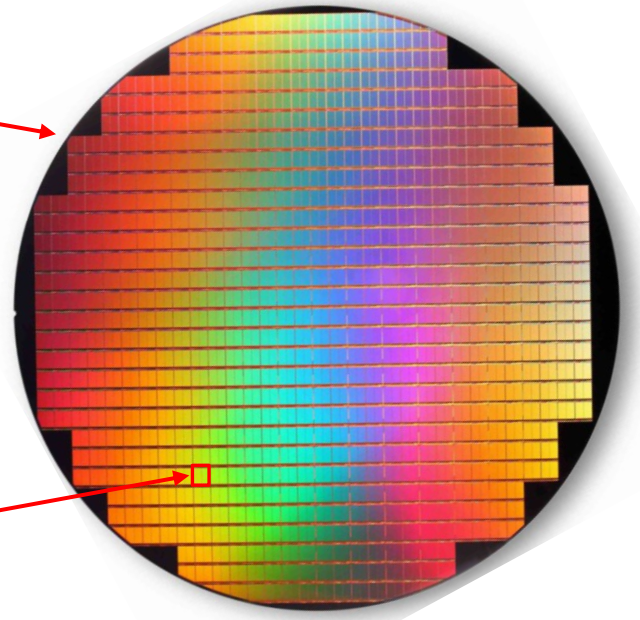
Wafer Defect Classifier

By Christopher Shaffer

What is a wafer?

A wafer is a thin slice of semiconductor (e.g., Silicon) which can contain microchips (“chips”/“dies”)





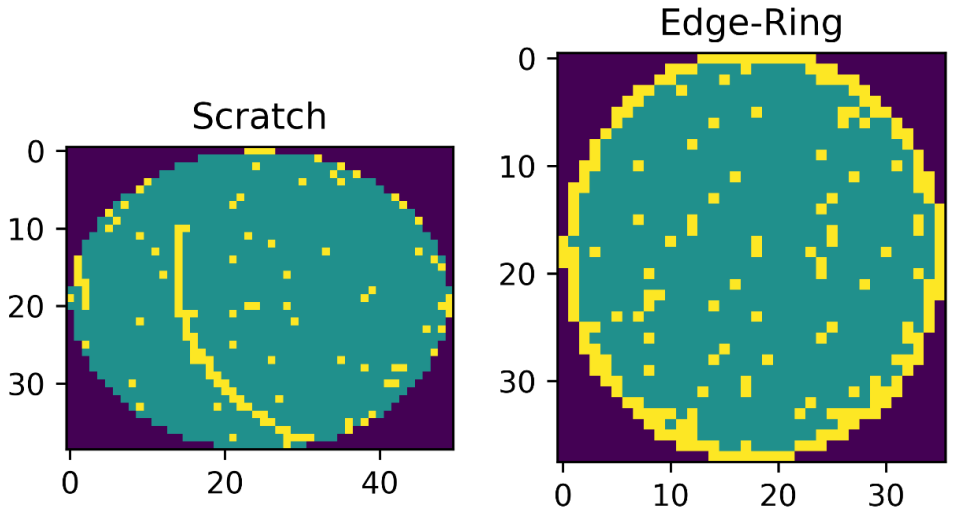
Motivation

- Semiconductor applications
  - Consumer electronics
  - Solar cells
  - Automotive/Industrial
- 2021 Forecast of semiconductor sales - $450 billion
- High Volume Manufacturing requires quality control
- Semiconductor inspection companies:
  - KLA-Tencor
  - Applied Materials
  - Hitachi

*https://www.semiconductors.org/wp-content/uploads/2020/06/2020-SIA-State-of-the-Industry-Report.pdf

Dataset & Goal

- ~800,000 real wafer maps from with various failure patterns
  - Electrical testing shows pass/fail map of all chips on wafer
  - Geometric patterns correspond to failure mechanisms
- 8 pattern classes



Wafer maps of 2 failure patterns

- Goal: Train a model to recognize failure patterns

Source: Public MIR-WM811K Corpus
http://mirlab.org/dataset/public/
https://www.kaggle.com/qingyi/wm811k-wafer-map

Data Cleaning

- Only 25,000 failed & labeled
  - Dropped all others
- Dropped extraneous columns
  - Die size, wafer index
- Categorically encoded labels
- Resized to uniform image size
- Pixel values to binary grayscale
- Formatted images for tensorflow

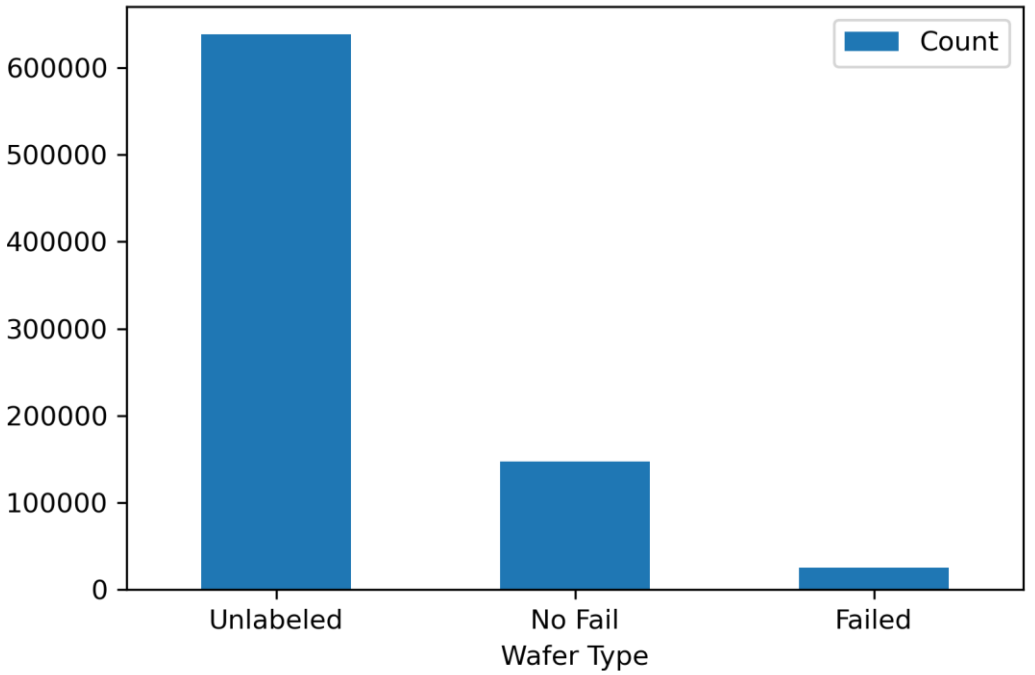
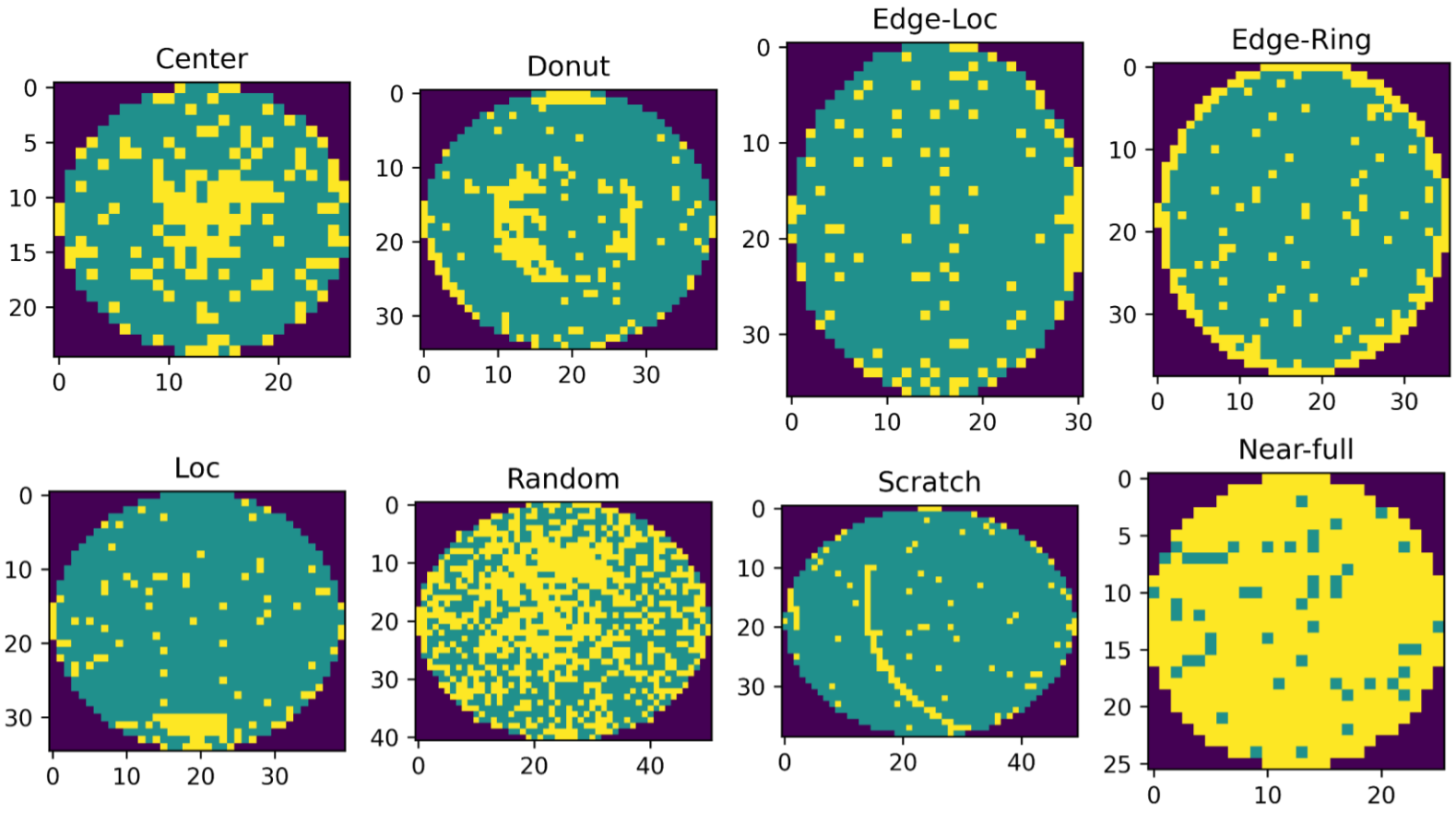


Image Classes and Dimensions

- Various aspect ratios and sizes
  - Roughly 10x10 to 100x100 pixels
- Resized all images uniform size
  - 24x24
  - 32x32
  - 48x48*

*Best performing and slowest

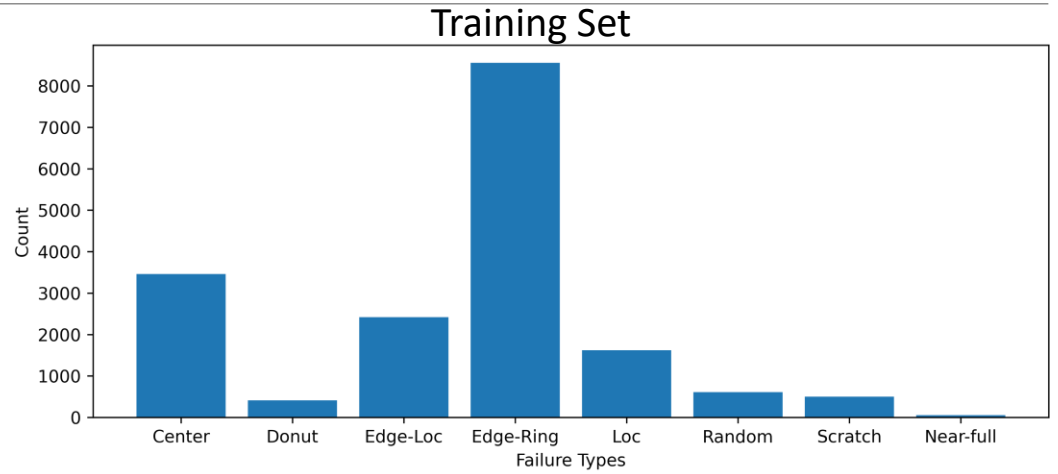


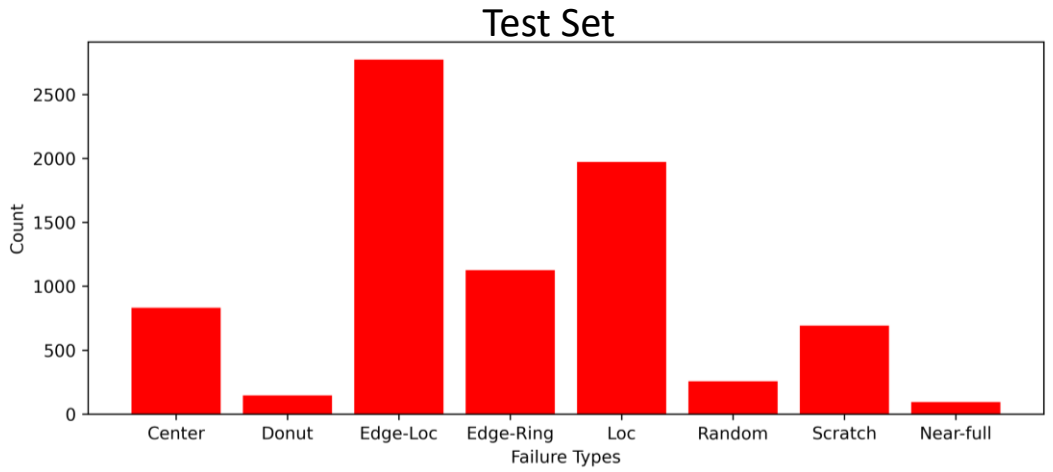
EDA

- Significant class imbalances
- Mismatch of class distributions between training and test set
  - Using train/test labels included in dataset
- Edge-Loc is plurality class for test data
  - 34% of test data

Class distribution (training)

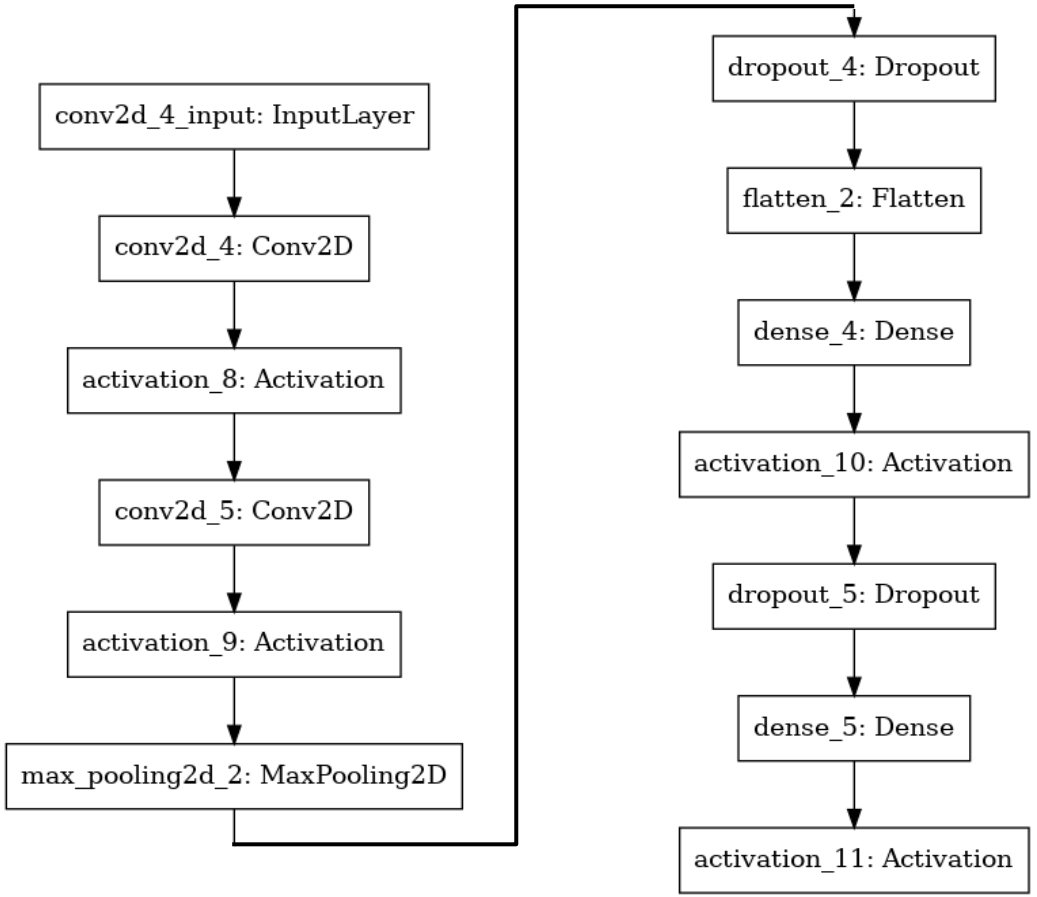
Class	Percentage
Center	19.6
Donut	2.3
Edge-Loc	13.7
Edge-Ring	48.5
Loc	9.2
Random	3.5
Scratch	2.8
Near-full	0.3



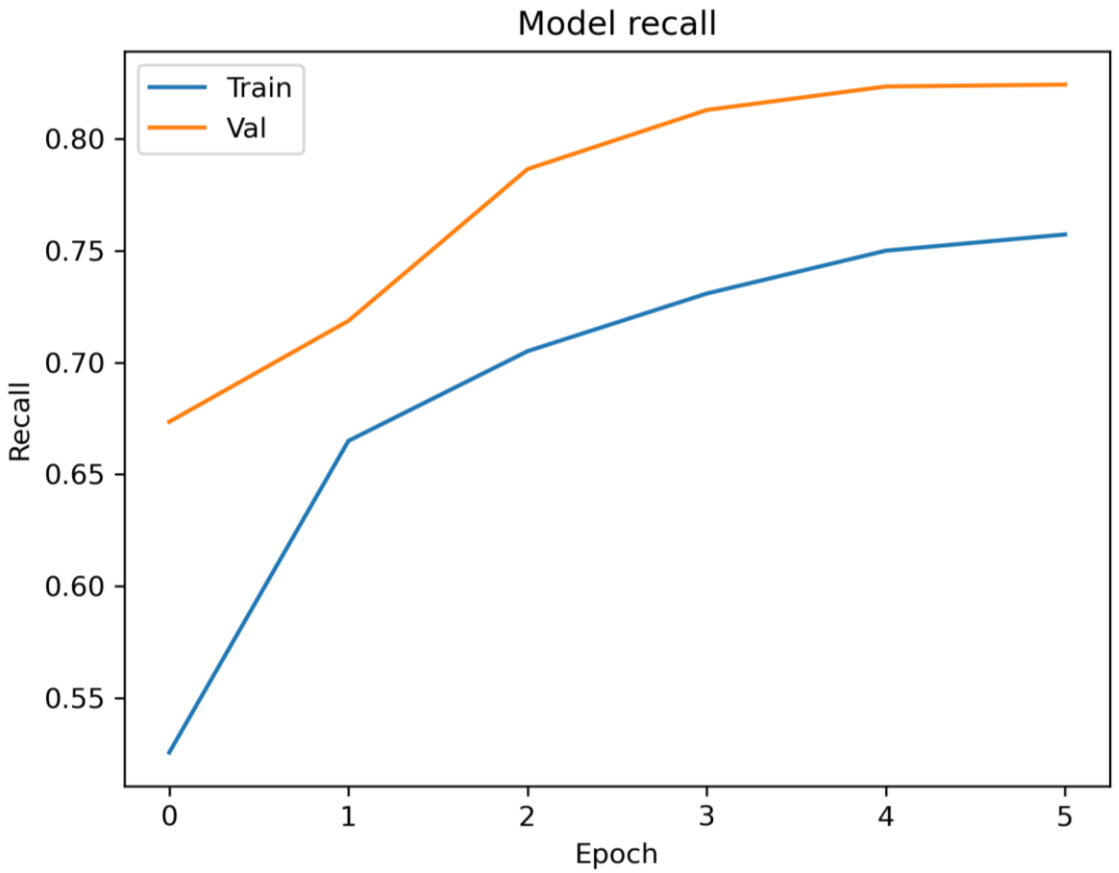


CNN Model

- Input dimensions: 48x48x1
- Dropout and pooling layers to prevent overfitting
- Final activation of softmax for multiclass classification



Final Results



Performance

	5-fold CV (avg)	Test Set
Recall	0.81	0.56
Precision	0.95	0.72
Accuracy	0.89	0.66*

*Compared to baseline of 0.34

Best Hyperparameters

Optimizer	Adam
Learning Rate	0.0003
# of Epochs	6
Batch Size	32
Image Size	(48,48)
Kernel Size	(5,5)

Appendix 1: SMOTE Oversampling

- Synthetic Minority Oversampling TEchnique
- Synthesizes new samples of minority classes using k-NNs
- All class counts set to ~8,000, plurality class count
- Results:

Without SMOTE	With SMOTE
Test recall: **0.46**	Test recall: **0.22**

Alternate approaches include:

- Variable class weights
- Augmented training data generators

Appendix 2: Testing Methodology

- Loss function - Categorical cross entropy
  - For multi-class classification
- Metric – **Recall** (aggregate over all classes)

Procedure

- **k-fold Cross-Validation**, k = 5
  - Only using training data
  - Hyperparameter tuning
- Training model over all training data
- Finally, evaluation over test set