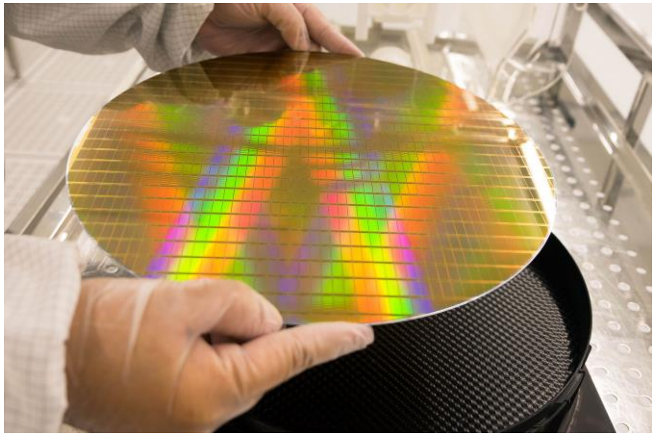
Ulster University

Semiconductor Defect Classification

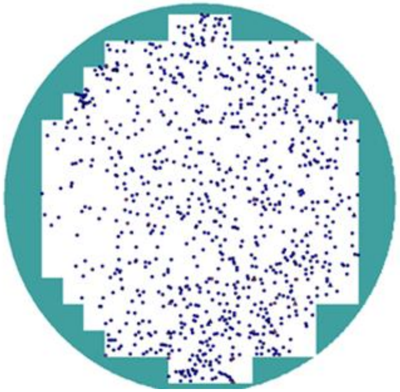
Terence Sweeney
Supervised by Professor. Sonya Coleman and Doctor. Dermot Kerr

Introduction

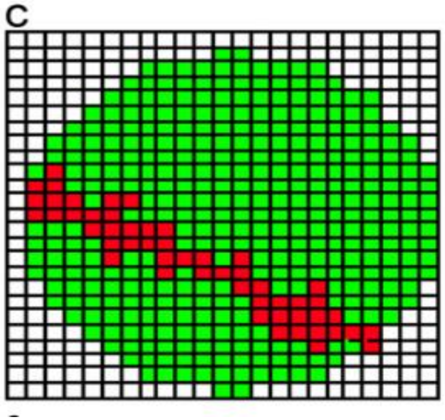
What is a semiconductor wafer?



A semiconductor wafer – Gep.com



Wafer Map – Rudolph

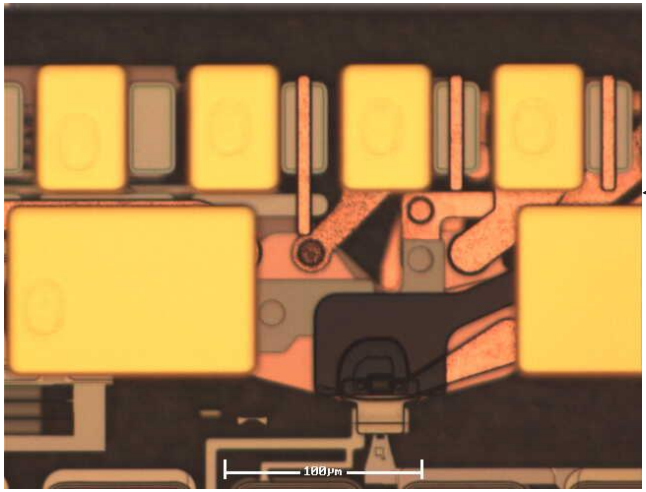


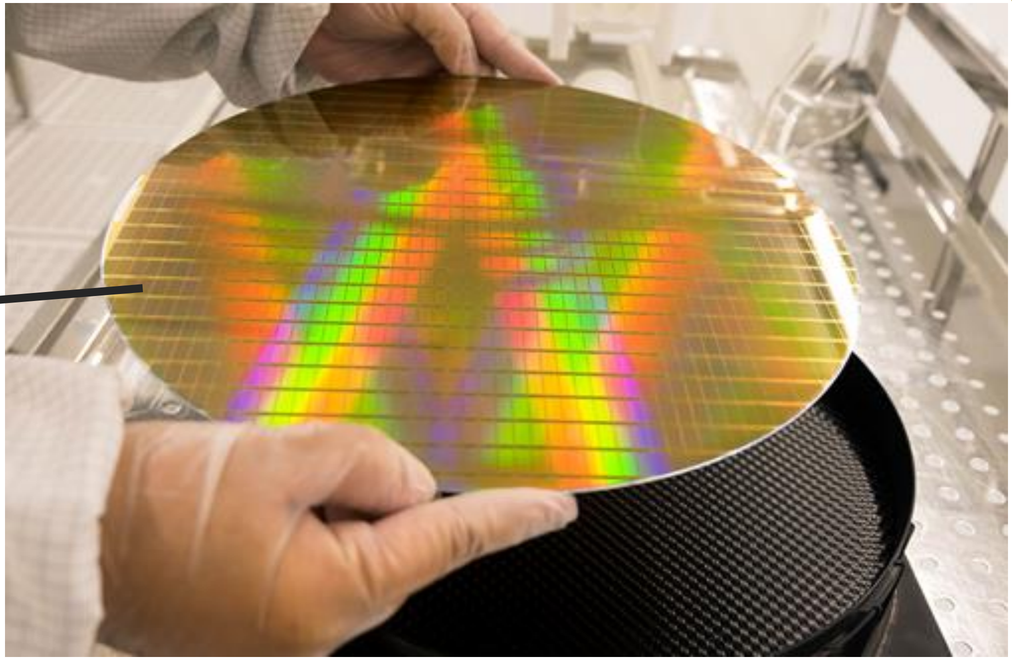
Line Defect –Ooi 2013

Introduction

Semiconductor wafer die

- We are focusing on the die image rather than the overall wafer.

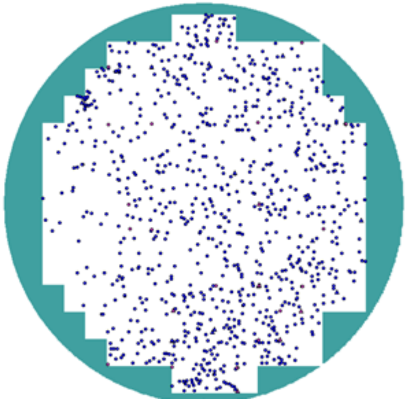


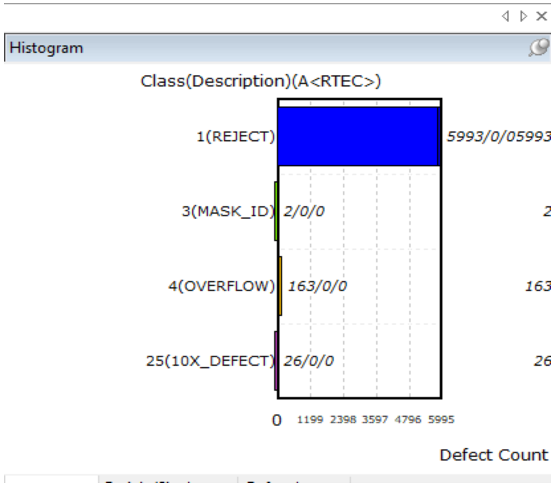


Current System

Rudolph NSX 105 and Discover

- Standard approach in industry : Golden image pixel-pixel comparison.
- Performed by specific machinery due to size.
- Seagate uses the Rudolph NSX-105 seen here along with the Discover software.



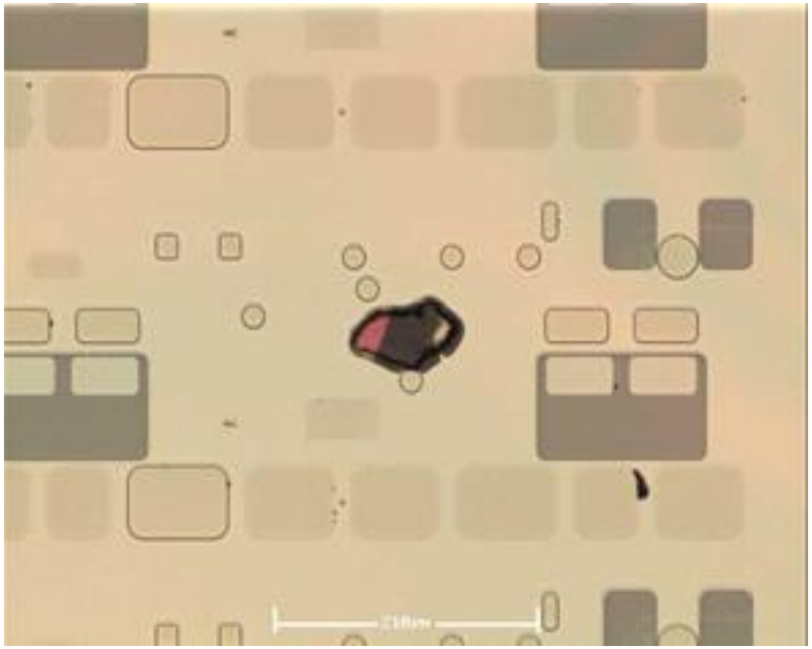


Ulster University

SEAGATE

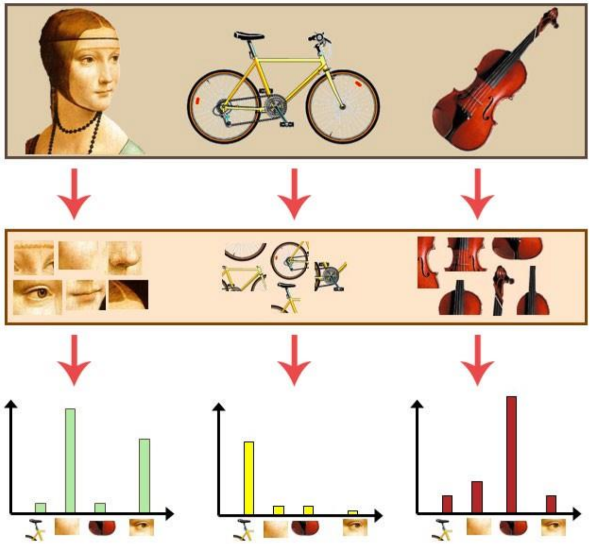
Defect classification

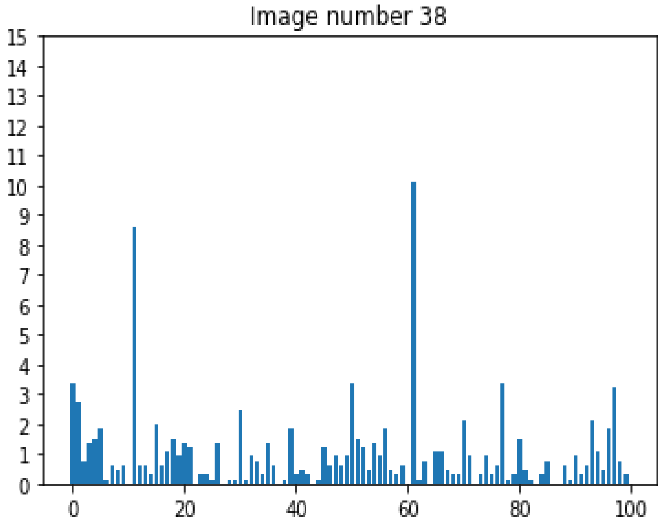
Defect Type
Splatter
Warping
Corrosion
Ripping
Scratching
Hole damage



Bag of Visual Words

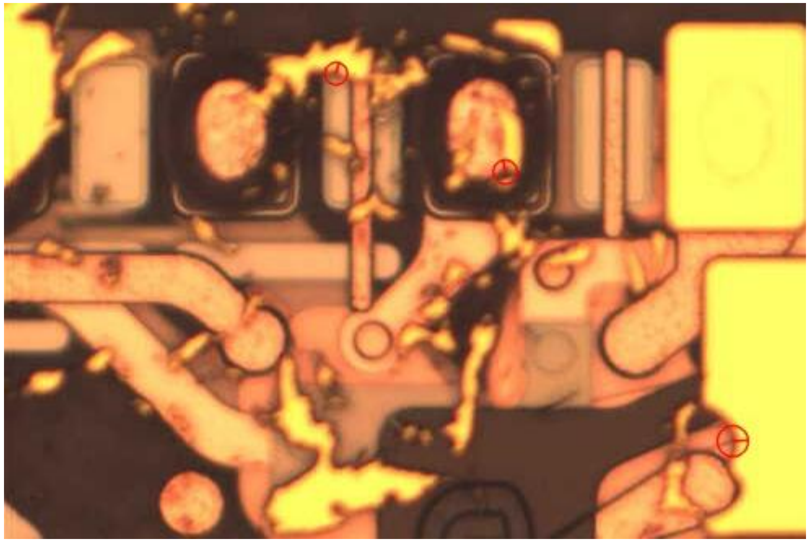
- Bag of words is a technique initially used for text.
- Bag of Visual words extends this into images with image features being the “visual words”.





BoVW – Custom Vocabulary

- A custom vocabulary is an augmentation of bag-of-visual-words, where the visual codebook that is created is augmented or pruned to focus on the features of most interest in the image



Ulster University

SEAGATE

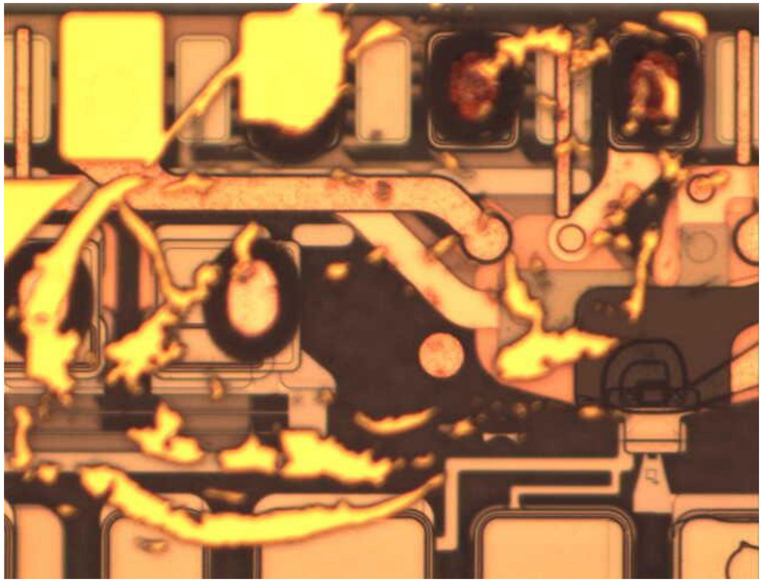
Experimentation overview

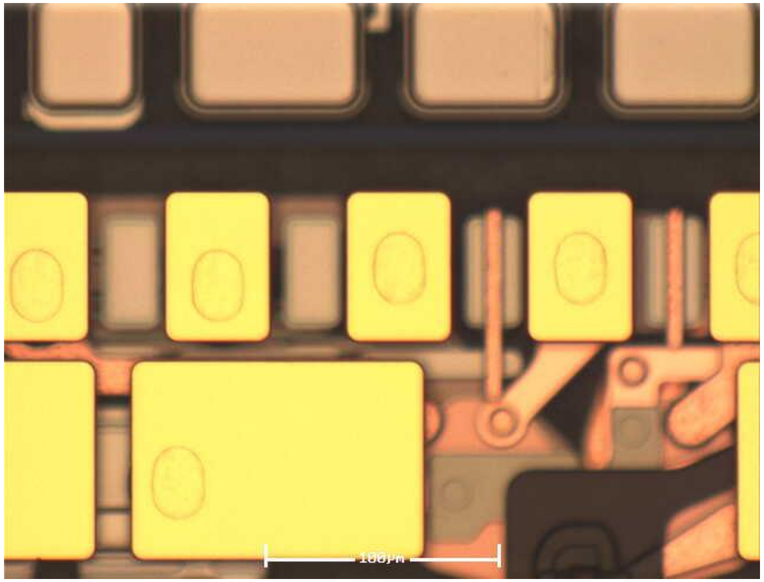
- The experiment for this paper was to evaluate the use of custom vocabulary in defecting semiconductor defects.

- To do this we initially created 1000 visual words from our training set in order to prune these down into defect only visual words.

Experimentation

- The methodology is applied to two classes, 100 Warp defect images and 100 control non-defect images
- Spilt into a 80/20 training/testing spilt.





Experimentation

- We ran the experiment and garnered the 1000 visual words from Warp and evaluated the results for this full set.

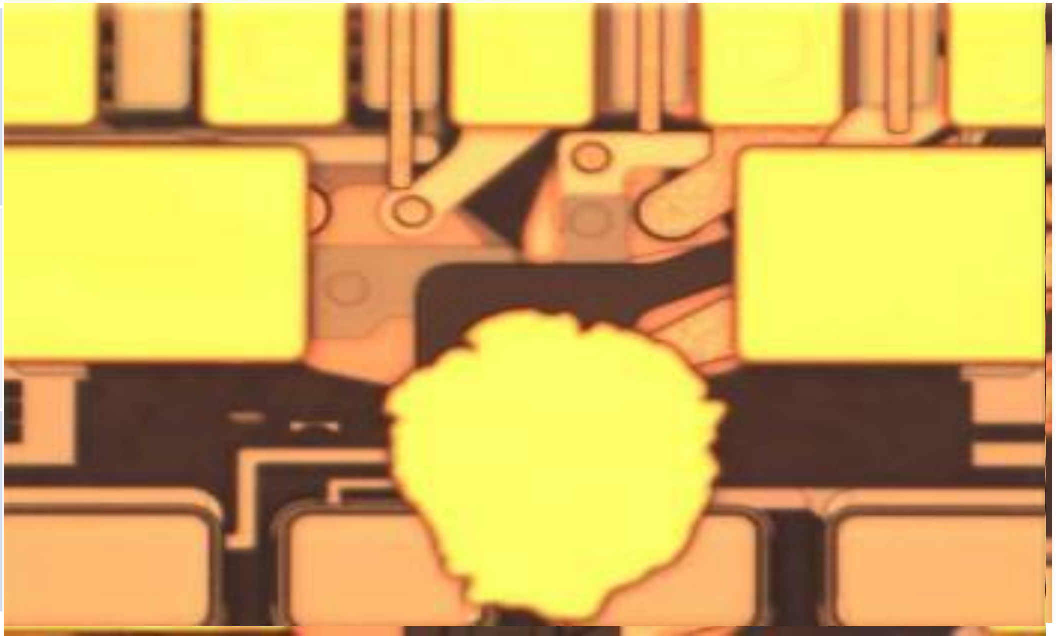
- After evaluation of the visual words, we retrieved 106 defect only visual words for the warp class and ran these through our ML classifiers to observe the results.

Computational speed Test	Training Time	Prediction Time	Highest Classification Accuracy
1000 Words	7m 37s	13s	100%
106 words	28s	9s	97%

	Accuracy 1000 visual words	Accuracy 106 visual words (custom vocabulary)
SVM – Poly C=1	**51%**	50%
SVM – Poly C=10	51%	**80%**
SVM – Poly C=100	56%	**97%**
SVM – RBF C=1	56%	**65%**
SVM – RBF C=10	**100%**	97%
SVM – RBF C=100	97%	**97%**

Subclass for defects

Defect Type	
Warp 1	
Warp 2	
Warp 3	



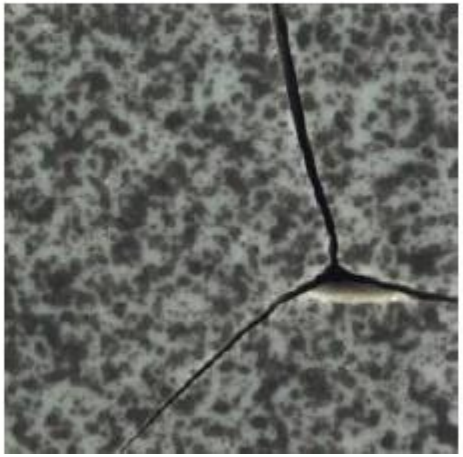
Subclass results:

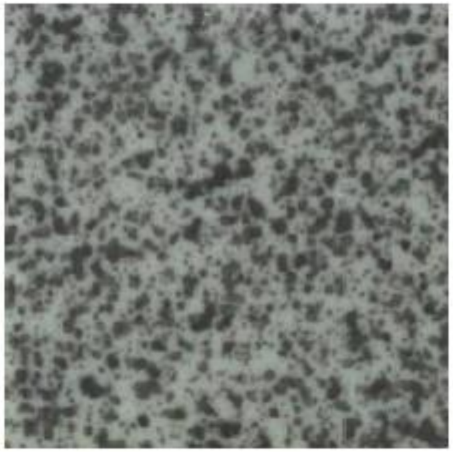
- The results gained from using these subclasses is below:

SVM			
Computational speed Test	Training Time	Prediction Time	Highest Classification Accuracy
1000 Words	7m 37s	13s	100%
106 words	28s	9s	97%
Subclass	**25s**	**6s**	**100%**
SVM PolyC-100	**100%**	97%	70%
SVM RBF C-1	87%	90%	70%
SVM RBF C-10	97%	**100%**	92%
SVM RBF C-100	**100%**	97%	95%

MVTEC Evaluation

- In order to further validate this system, we use the MVTEC anomaly detection dataset
- From the dataset we selected the Tile Crack image set which contains 20 images, 10 for training and 10 for testing, along with a control class, again using 10 for training and 10 for testing





Ulster University

SEAGATE

MVTEC Evaluation

- In line with the previous experiment, as the SVM performed best, we use only an SVM with all 1000 visual words and the defect only visual words, for which 69 were detected for this dataset.

SVM		
	Full 1000 Words	**69 Defect only Words (Custom Vocabulary)**
SVM Linear C-1	72%	80%
SVM Linear C-10	72%	80%
SVM Linear C-100	72%	80%
SVM Poly C-1	72%	80%
SVM Poly C-10	72%	80%
SVM Poly C-100	72%	**97%**
SVM RBF C-1	72%	80%
SVM RBF C-10	72%	80%
SVM RBF C-100	82%	**97%**

Ulster University

SEAGATE

Conclusions

- We have presented an approach to semi-conductor wafer defect classification by utilizing the bag of visual words method with a custom vocabulary formed from a reduced set of visual words.

- We have demonstrated that this novel approach achieves competitive accuracies when compared with the use of a larger set of visual words (1000) but is much more computationally efficient as demonstrated by the presented run-times.

Ulster University

SEAGATE

Questions?

Thank you for Listening

Ulster University