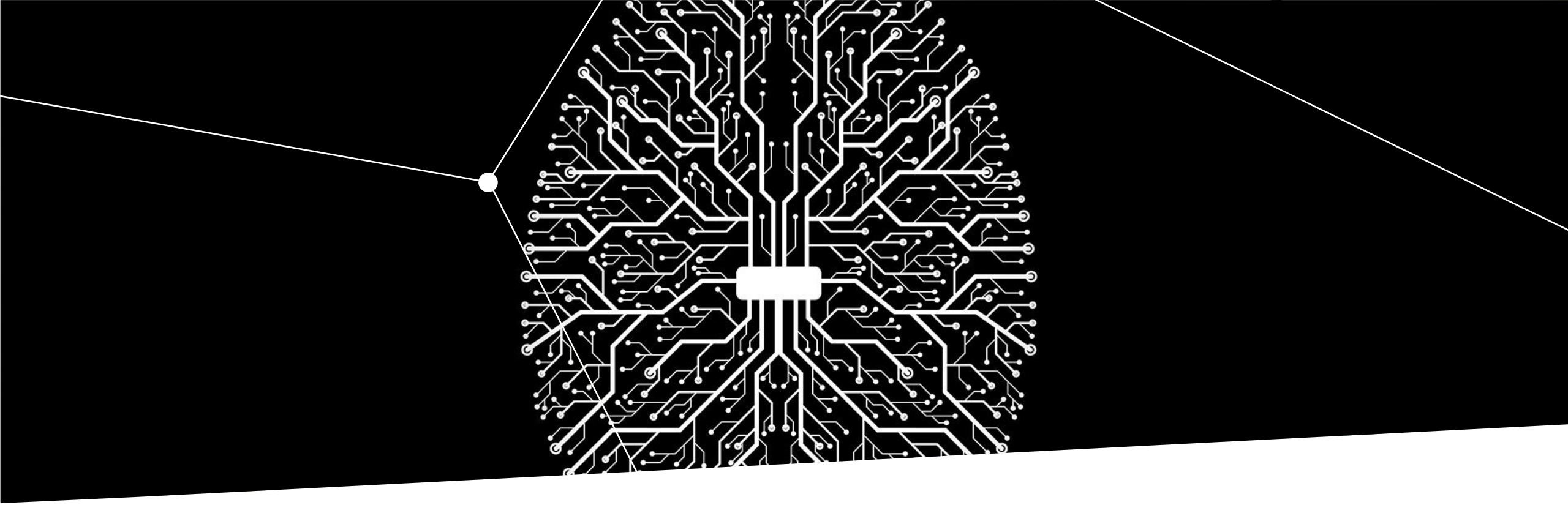


Using Image Clustering for Auto Annotation of Defect ICs

Zhu Ye (Juliana)
2023/05/25

Background Introduction

VOILA (Virtual Optical Image cLAssification)

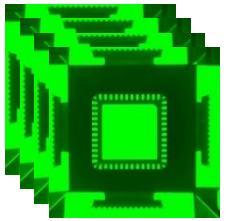
- Automated-optical-inspection (AOI) tools are deployed prior to shipment of ICs to ensure **out-going quality**
- Equipment relies on rule-based algorithms that have **low classification accuracy**
- An evaluation showed 50% of images were either **misclassified** or **false failures**
- High volume of images (~100k) generated daily

ICs

Automated-optical-inspection (AOI) machine

Rule-based algorithms

Reject images



Engineering Intervention:
optimize recipe
re-scan overkill units

VOILA

Image Classifier

A solution based on **Deep Learning approach** has been developed and deployed to **improve** the optical inspection process:

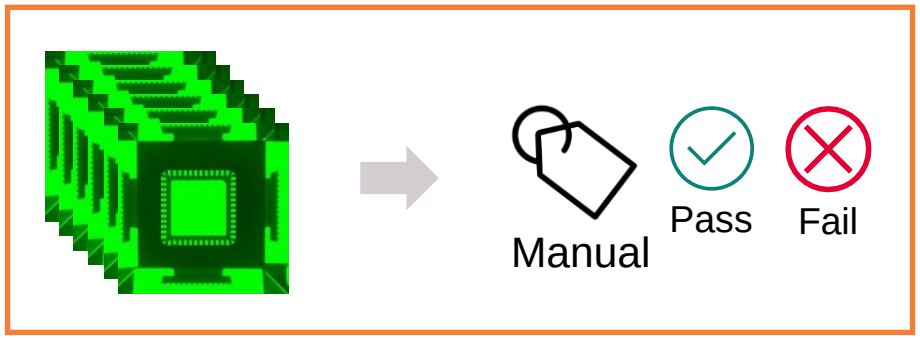
- **Recover** false failures at point of process
- **Classify** images accurately based on defined failure modes
- Alert and **trigger timely actions** for lots with high overkill

Business Value: Improve yield by reducing the number of over-rejections

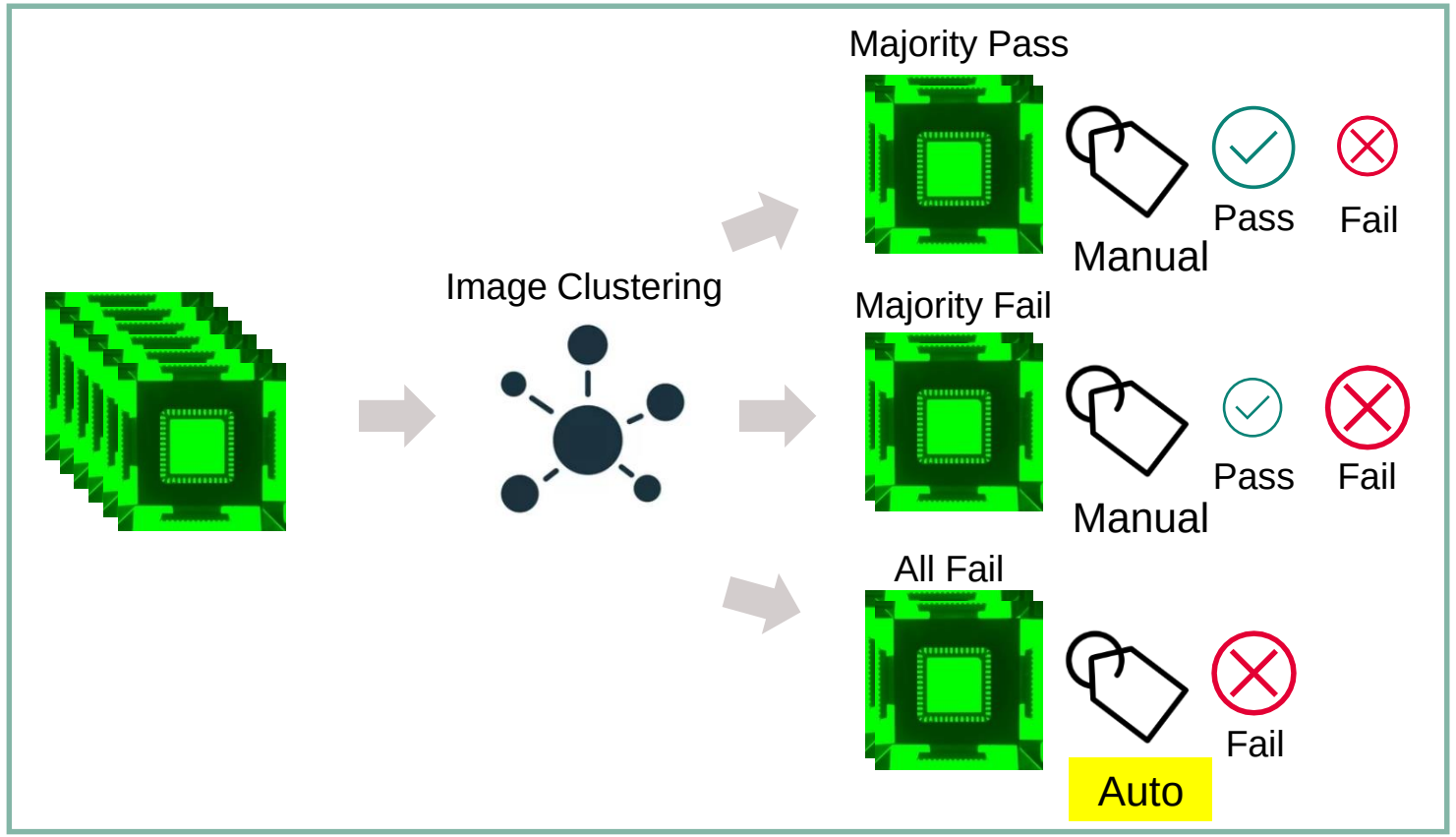
- VOILA harvesting on existing packages expected €1.35mio savings by end of FY2223
- Potential to harvest even more with scope extension planned on more packages and defects as well as different AOI machine models

Problem Statement and Objective

To accommodate the scaling up of the Deep Learning model, including additional packages, defects, and different AOI machine models, an increased amount of labelled data is required for training. However, so far all the labelling was done manually by our domain experts, which is time consuming.



Using image clustering for auto annotation to reduce the time for manual labelling.



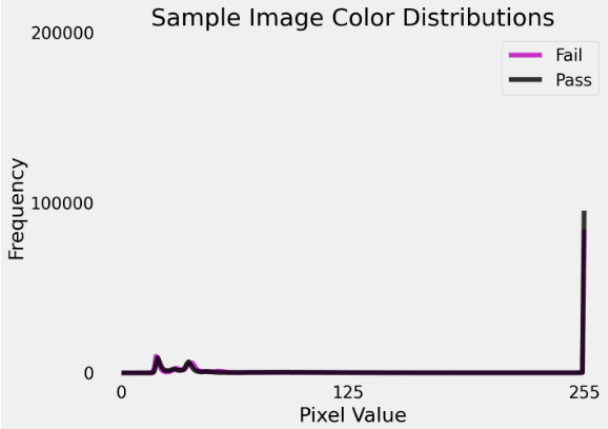
Exploratory Data Analysis (EDA)

Dataset

Company in-house data: 3892 KLA 4-side defect bright mold images of VQ48 for initial trial (binary class pass/fail)
KLA: AOI machine model
4-side defect bright mold: defect mode
VQ48: IC package name

Image Class Distribution

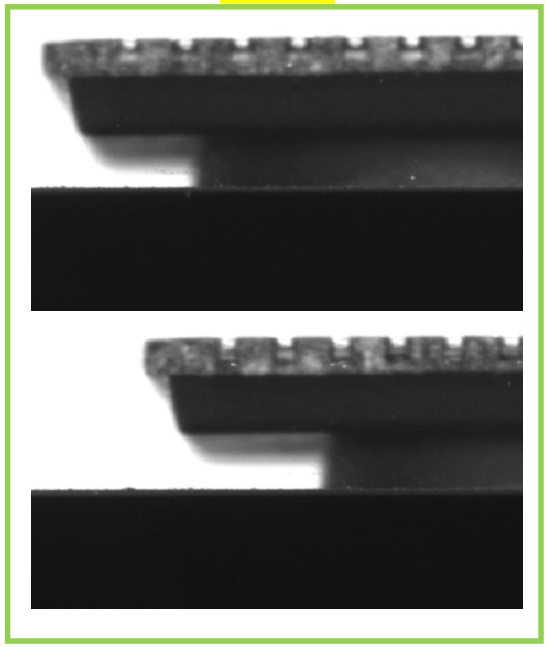
Image Colour Distribution



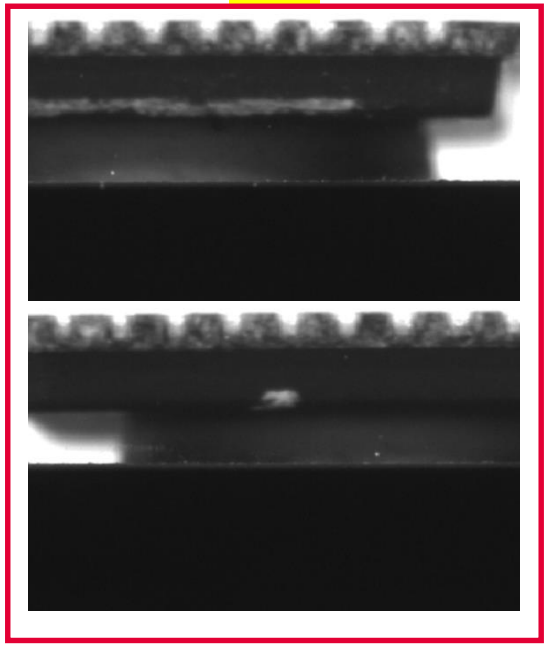
- Unbalanced data, more pass (61%) than fail (39%)
- Same colour distribution for pass and fail images

Sample Images

PASS

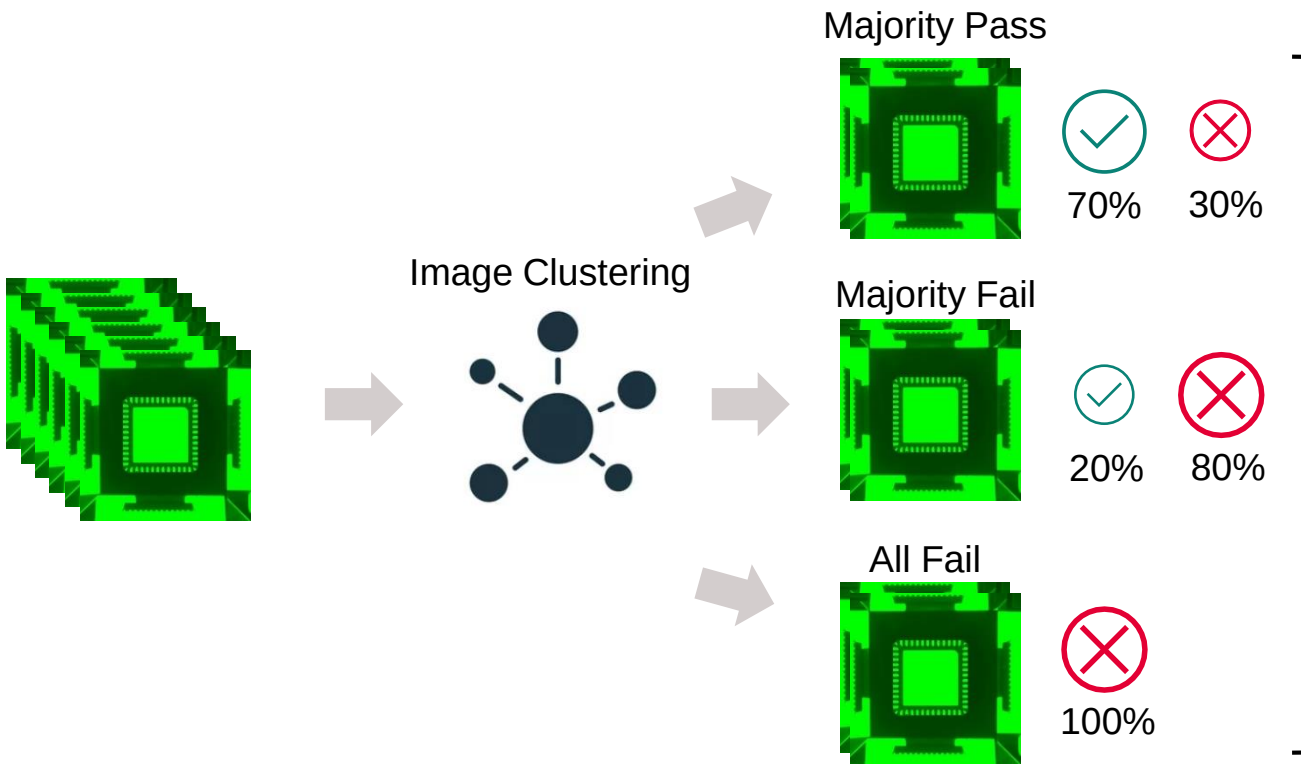


FAIL



- Pass: without white stain
- Fail: with white stain (line/dot)

Success Metrics for Evaluation

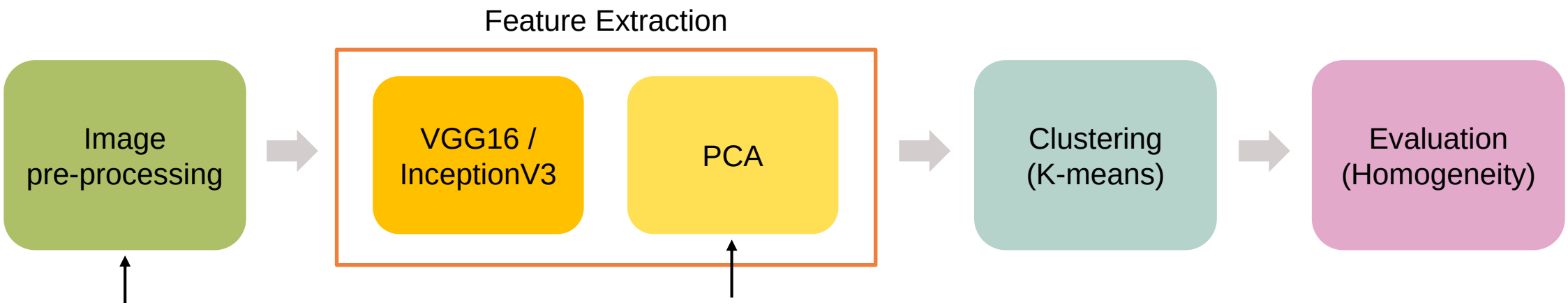


$$\text{(Weighted) Average Cluster Homogeneity} = \frac{\sum cluster\ size\ * majority\%}{Total\ sample\ size}$$

- **Homogeneity**: Using ground truth labels (pass/fail) to calculate how many % of majority in each cluster (somewhat like “precision”)
- Ideal case is 100% pass (or 100% fail) in every cluster, worst case is 50%-50% pass and fail mixed in every cluster

- Calculate **Average Cluster Homogeneity** as overall score (value range is 0.5 -1.0)

Modelling Methods

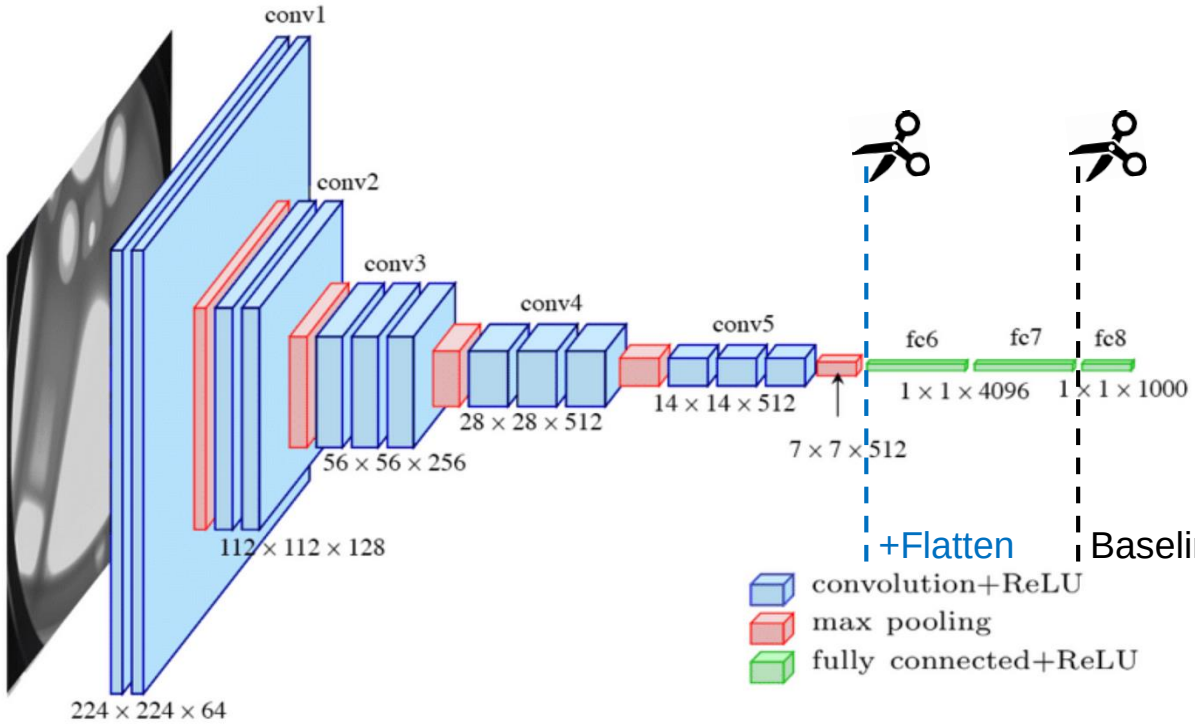


- No colour scale adjusting
- Resize to 224 x 224 before modelling

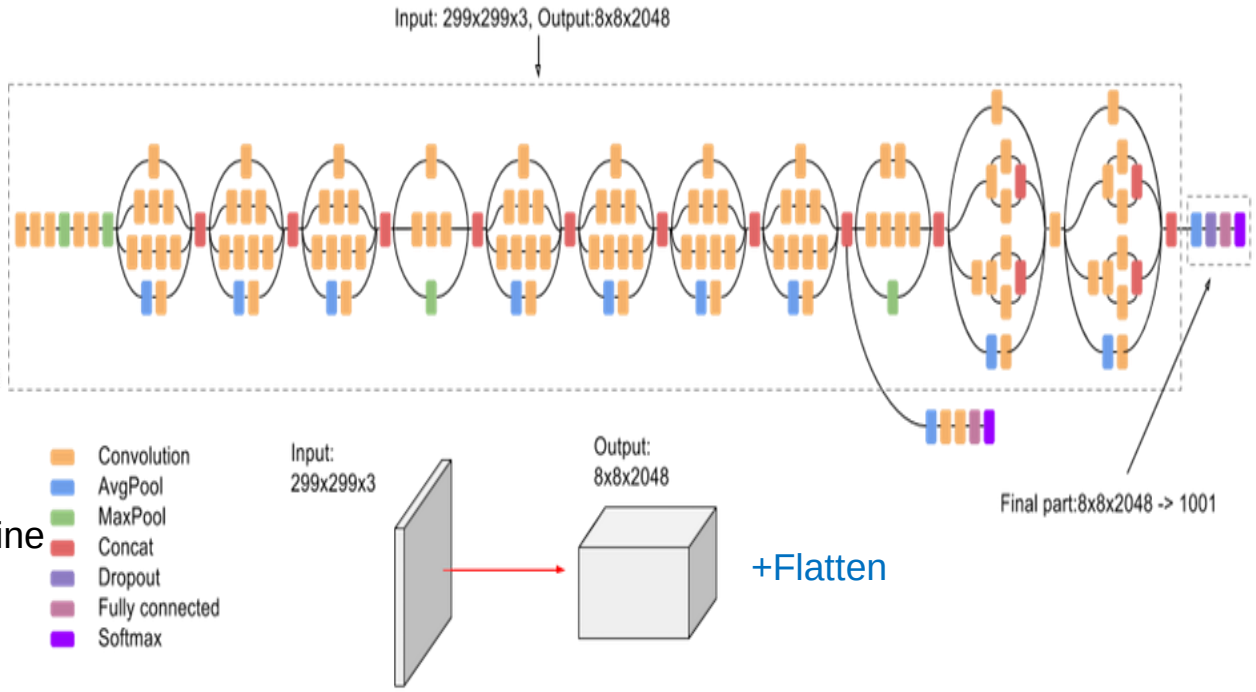
- Select key components no. when Cumulative Explained Variance >0.95

Tailor and Transfer VGG16 & InceptionV3 for Feature Extraction

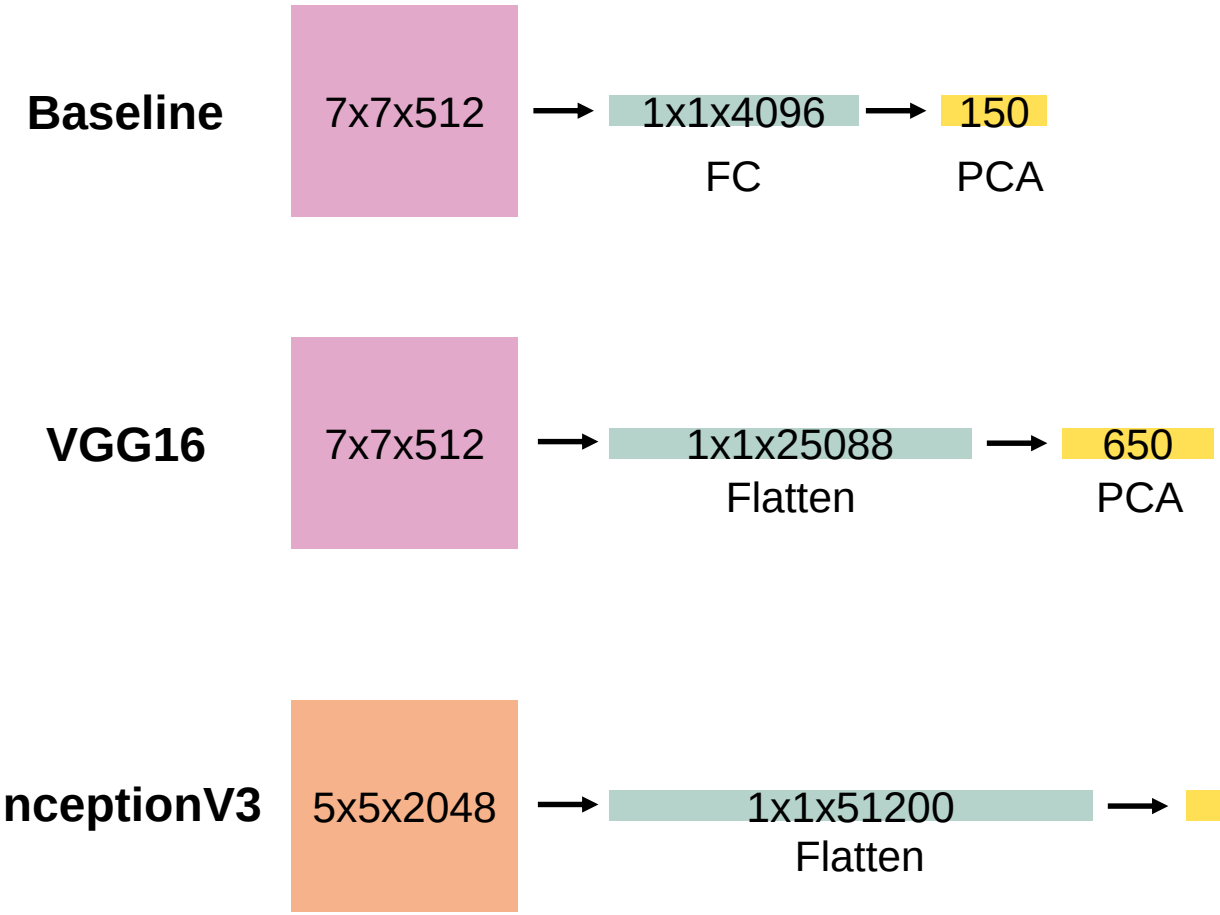
VGG16 Architecture



InceptionV3 Architecture

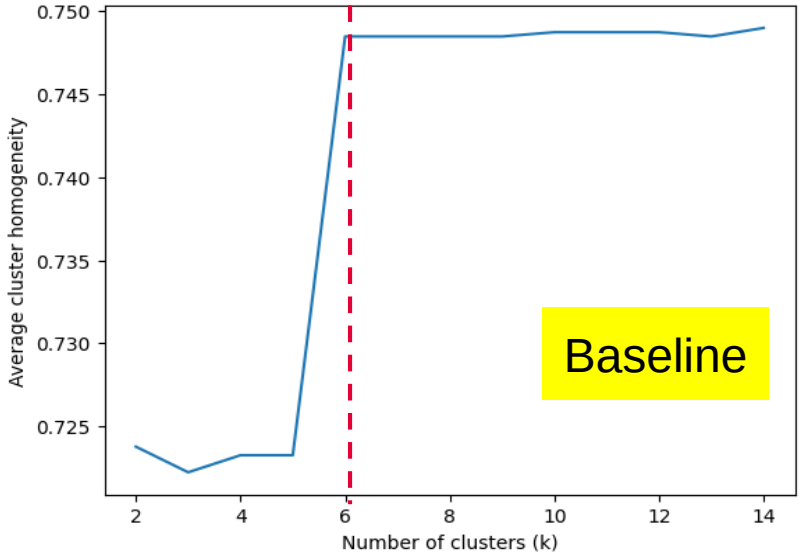


Feature Extraction Results of 3 Methods



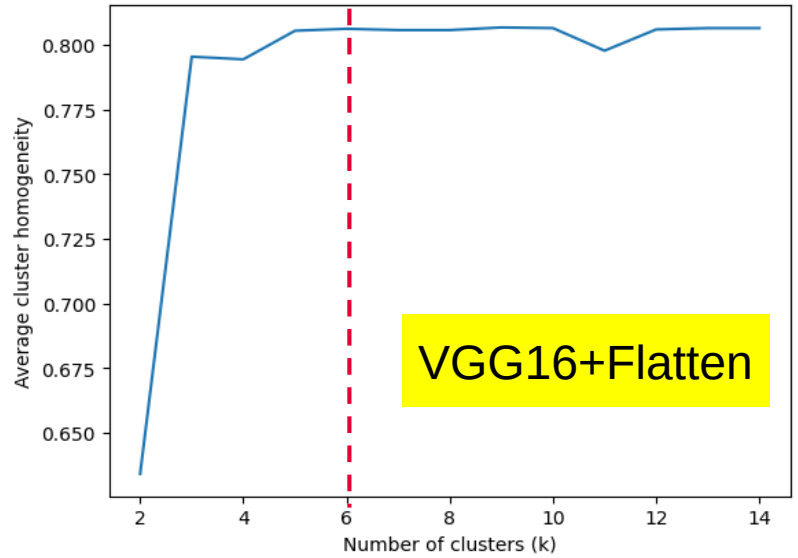
- Increasing image vectors after flatten will result in the increase of key components no. of PCA, which can improve the clustering results
- InceptionV3 extracted the largest numbers of features among 3 methods

Modelling Results: VGG16+Flatten vs. Baseline



Baseline	Cluster 2	Cluster 0	Cluster 4	Cluster 1	Cluster 5	Cluster 3
Cluster size	706	142	202	257	1379	1206
Majority class	Pass	Fail	Fail	Fail	Pass	Pass
Homogeneity (%)	74.8%	84.5%	96.0%	100%	71.9%	68.2%

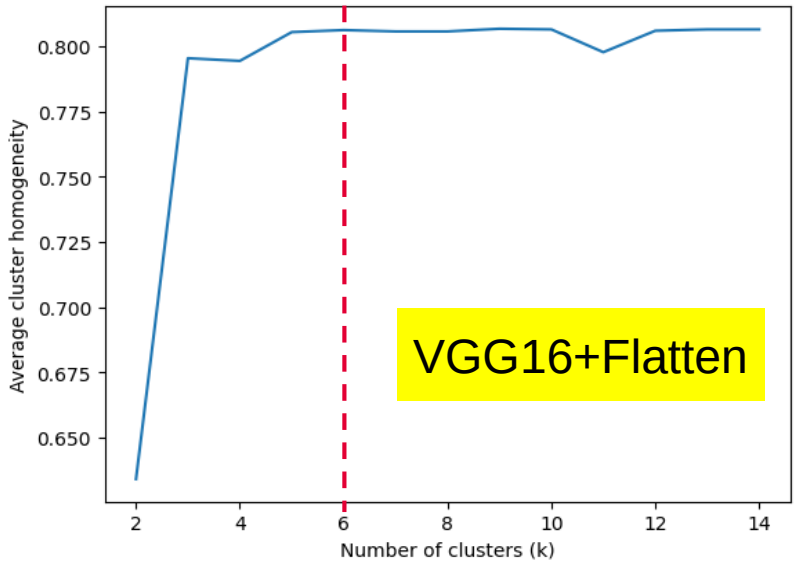
› K=6 has Average Cluster Homogeneity **0.748** and the homogeneity of each cluster is all above **68**%



VGG16+Flatten	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Cluster size	1078	365	202	257	611	1379
Majority class	Pass	Fail	Fail	Fail	Pass	Pass
Homogeneity (%)	81.8% ↑	94.2% ↑	96.0%	100%	74.1% ↑	73.1% ↑

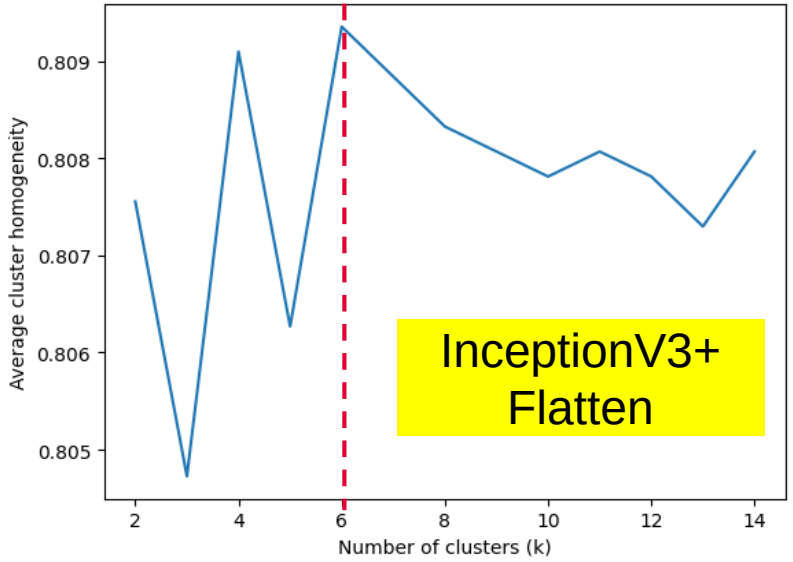
› K=6 has Average Cluster Homogeneity **0.806** and the homogeneity of each cluster is all above **73**%, better than baseline results

Modelling Results: VGG16+Flatten vs. InceptionV3+Flatten



VGG16+Flatten	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Cluster size	1078	365	202	257	611	1379
Majority class	Pass	Fail	Fail	Fail	Pass	Pass
Homogeneity (%)	81.8%	94.2%	96.0%	100%	74.1%	73.1%

› K=6 has Average Cluster Homogeneity **0.806** and the homogeneity of each cluster is all above **73**%

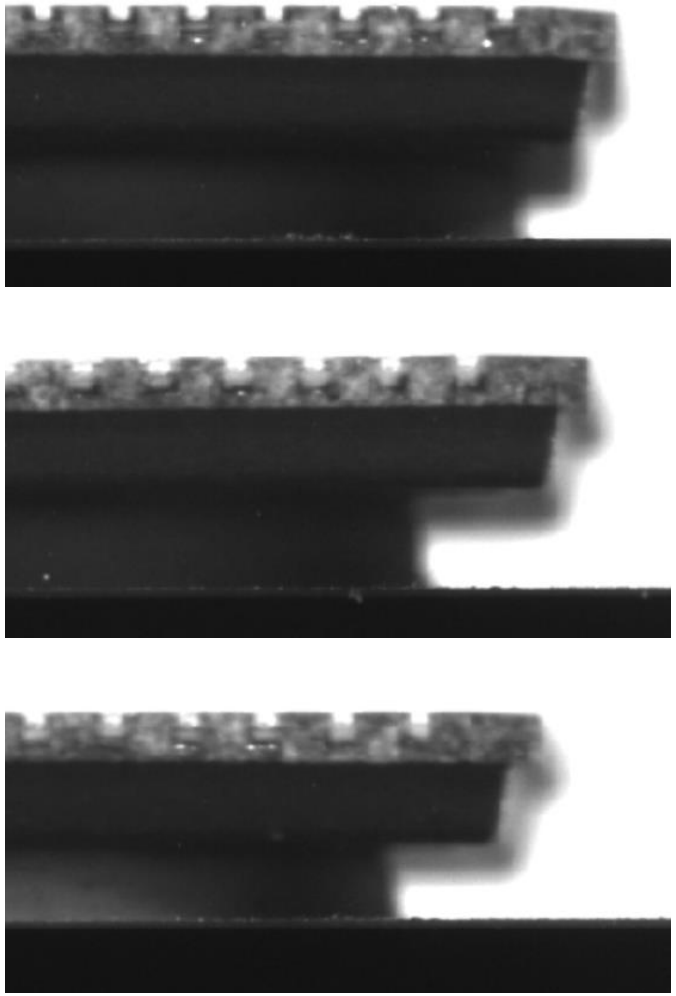


Incept + Flatten	Cluster 0	Cluster 4	Cluster 5	Cluster 2	Cluster 3	Cluster 1
Cluster size	1698	200	352	260	777	605
Majority class	Pass	Fail	Fail	Fail	Pass	Pass
Homogeneity (%)	79.3% ↓	96.0% ↑	97.4% ↑	100%	76.4% ↑	68.6% ↓

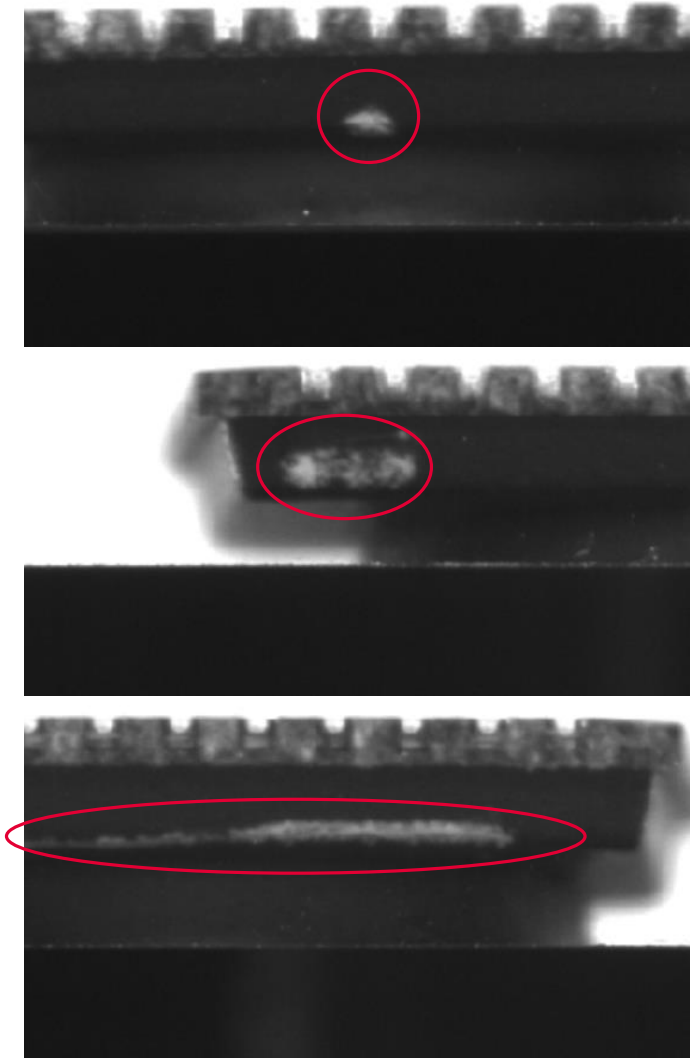
› K=6 has Average Cluster Homogeneity **0.809** and the homogeneity of each cluster is all above **68**%, comparable with VGG16 + Flatten results

Representative images of majority for each cluster (VGG16+Flatten)

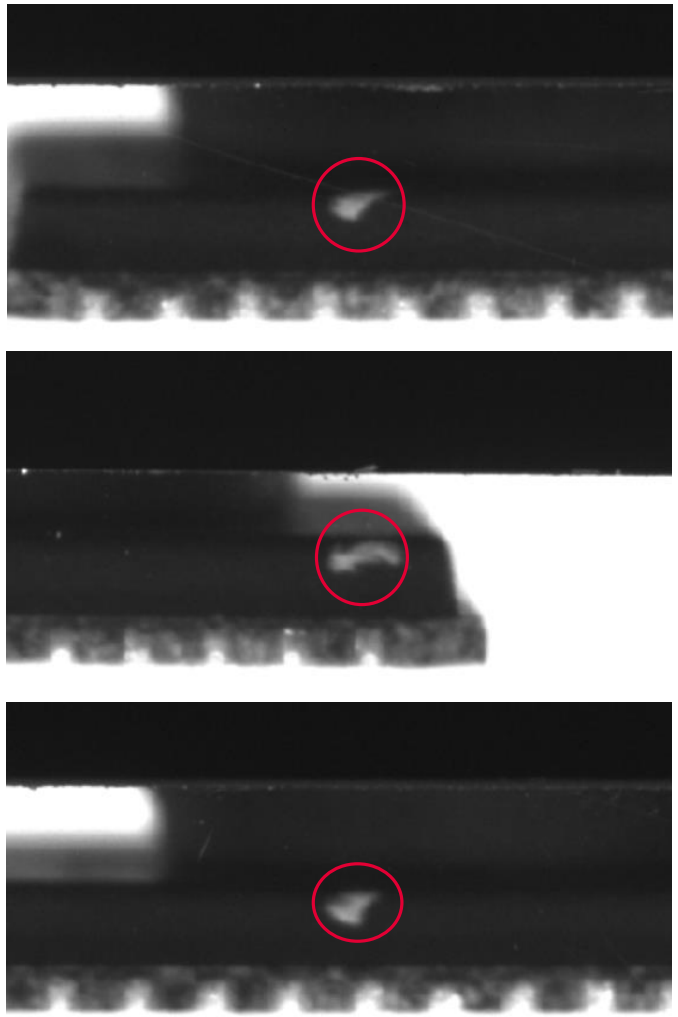
› Cluster 0 (right up, 81.8% pass, no scratch/stain)



› Cluster 1 (up, 94.2% fail, line/dot stain)

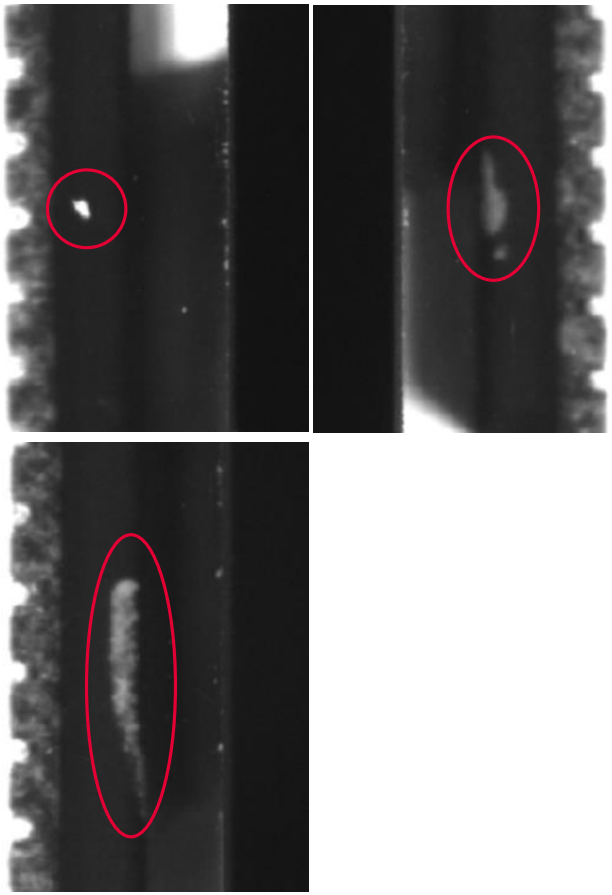


› Cluster 2 (down, 96.0% fail, dot stain)

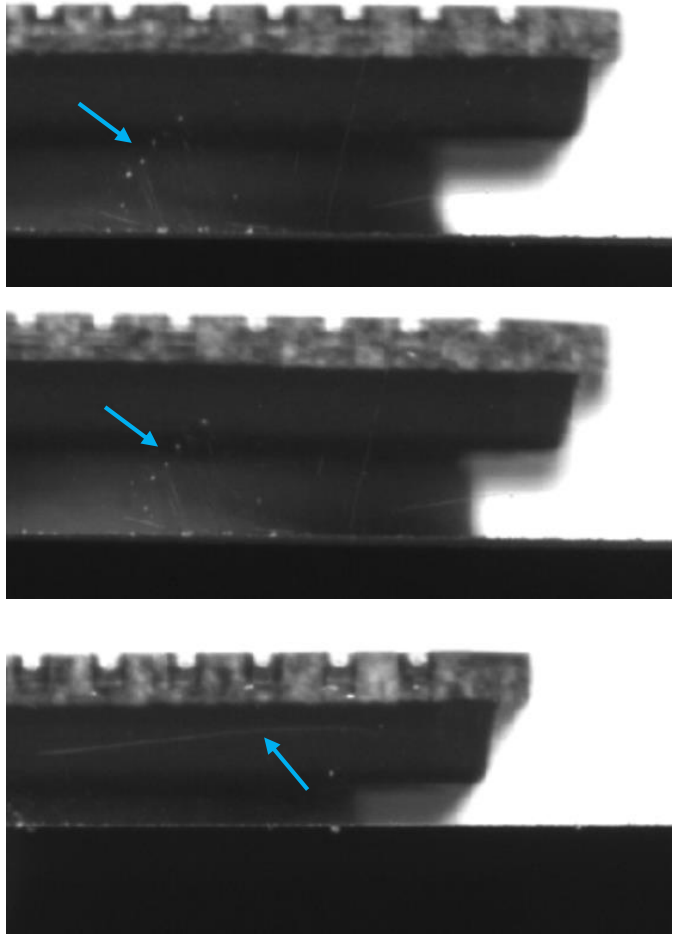


Representative images of majority for each cluster (VGG16+Flatten)

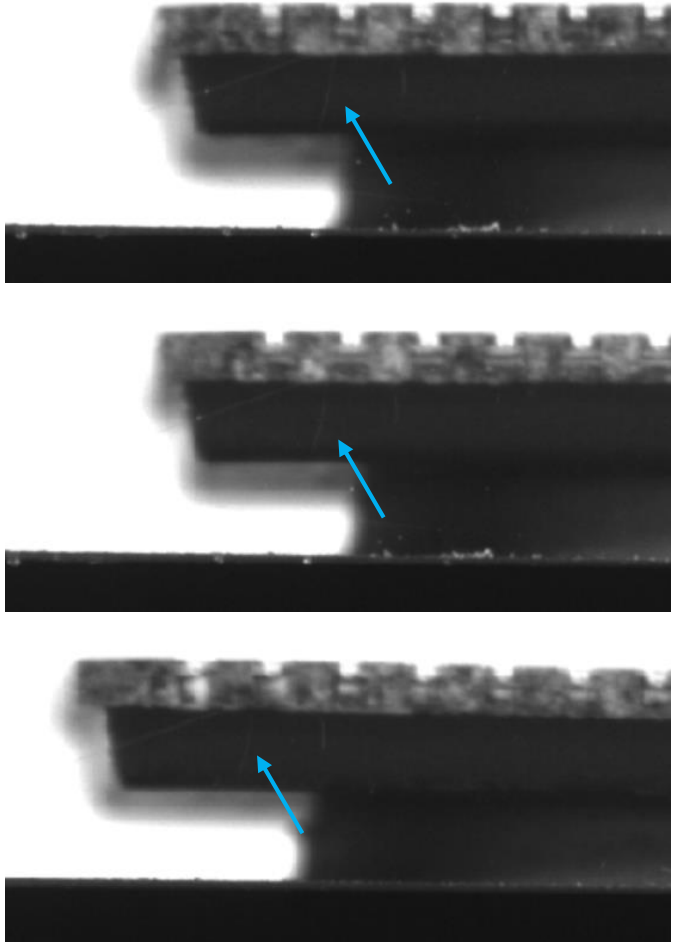
- Cluster 3 (left/right, 100% fail, line/dot stain)



- Cluster 4 (right up, 74.1% pass, small stain points and scratch lines)



- Cluster 5 (left up, 73.1% pass, scratch lines)



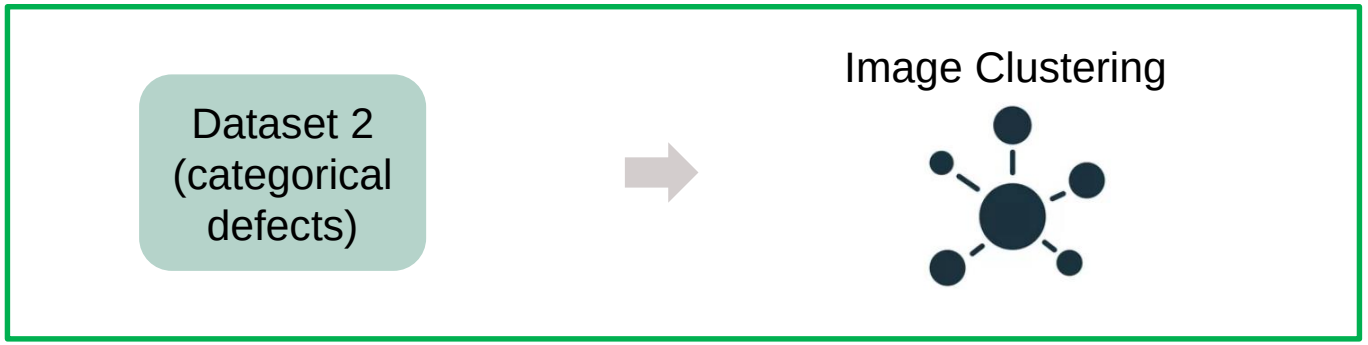
Conclusions and Recommendations

- 3 methods of transfer learning for feature extraction of image clustering were applied:
  - Removed the last FC layer of VGG16 as baseline
  - Removed the last 3 FC layers of VGG16 + flatten
  - InceptionV3 + flatten

- Using VGG16+Flatten for feature extraction and k-means with k=6 can have an Average Cluster Homogeneity **0.806** and the homogeneity of each cluster all above **73**%, which can help to reduce the time of manual labelling

- The model can be used for clustering of similar IC images with the expectation of similar homogeneity

Next Step



- Dataset 2 will have different defect categories for the failures
  - Cluster all data with pass and fail with different defect categories
- Use in-house VOILA Offline trained classification model for transfer learning of clustering
- Use other clustering method (e.g. DBSCAN)

Part of your life. Part of tomorrow.