VISION 23 Challenge Track 2 - Data Generation for Defect Detection

Davy, ShiYu Long, Han Xu, Jun Wang, Xiao Xiao

Focusight

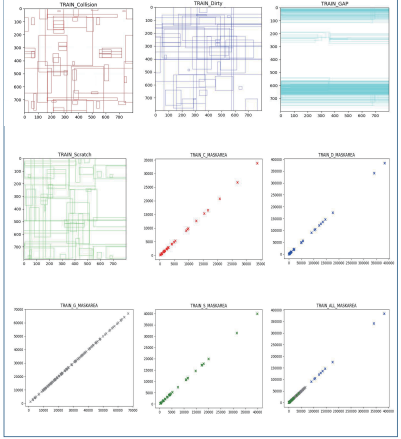
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Analysis

- **The distribution of the original training set**

Type	NG				OK
	Collision	Dirty	Gap	Scratch	
Quantity	51	43	194	53	1885
Percentage	2.29%	1.93%	8.72%	2.38%	84.68%

- **The location characterstic of defects in train dataset**

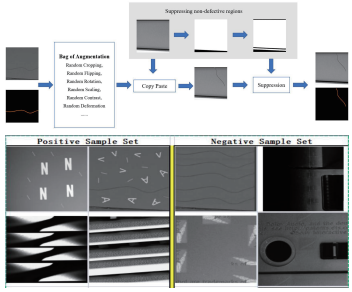


Methods

- **Step 1: Data Augmentation**
  - Random small rotation.
  - Random integer multiples of 90° rotation.
  - Random flip.
  - Random brightness.
  - Some of the different samples are cleaned.
  - Randomly expand to 11000.

The final score of this validation is : 0.216.

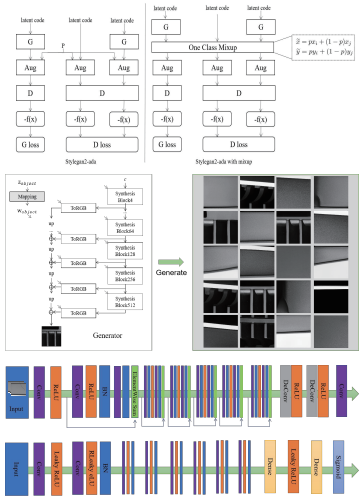
- **Step 2: Defect Mask-CopyPaste**



- Augmenting positive samples with excessive killing features.
- Conducting appearance transfer on certain negative samples using DM-CP.

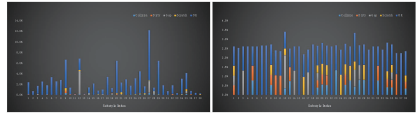
The final score here is : 0.248.

- **Step 3: GAN**



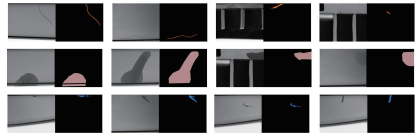
The richness of defect appearance was increased based on GAN method, and we achieve 0.256 score on leader board.

- **Step 4: Subtype Analysis**



There is a serious imbalance between the sub-patterns of positive and negative samples. Through the analysis, the sample balance is carried out based on sub-patterns. The score is 0.263 after applying sample balance.

- **Step 5: Ensemble**



We achive the final score 0.270.

Experimental Result

- **The mAP results of Differenet Methods**

Method	mAP
Baseline	0.172
Data Augmentation	0.216
Local Data Augmentation + DM-CP	0.248
Local Data Augmentation + GAN	0.256
Data Augmentation + Sub-Appearance	0.263
Data Augmentation + Sub-Appearance + GAN + DM-CP	0.270