EDD-Net: An Efficient Defect Detection Network

Tianyu Guo, Linlin Zhang, Runwei Ding, Ge Yang

Key Laboratory of Machine Perception

Shenzhen Graduate School, Peking University, China

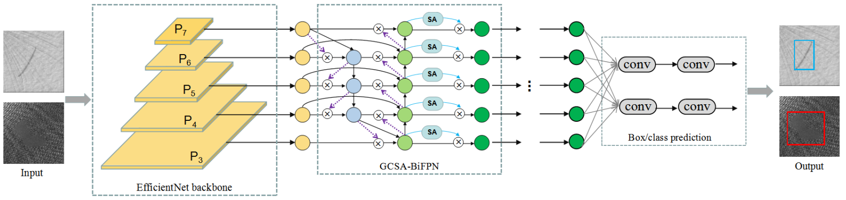
Introduction

As the most commonly used communication tool, the mobile phone has become an indispensable part of our daily life. The surface of the mobile phone as the main window of human-phone interaction directly affects the user experience.

It is necessary to detect surface defects on the production line in order to ensure the high quality of the mobile phone. However, the existing mobile phone surface defect detection is mainly done manually. Currently, there are few automatic defect detection methods to replace human eyes. How to quickly and accurately detect the surface defects of the mobile phone is an urgent problem to be solved.

Pipeline of EDD-Net

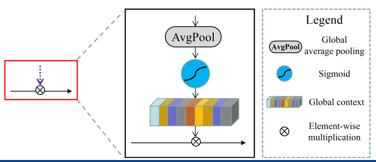
The input image first passes through a lightweight backbone network EfficientNet to extract features. Then, the GCSA-BiFPN which consists of the bidirectional feature pyramid module, global context module, and spatial attention module is proposed to obtain more discriminative features. Finally, a box/class prediction network is used to achieve defect detection.



Global Context Module

In BiFPN, the global context information can be obtained by the following:

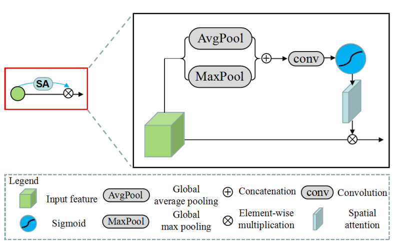
$$GC(F) = \sigma(AvgPool(F))$$



Spatial Attention Module

The spatial attention module is connected after each pyramid layer. The spatial attention can be obtained by the following equation:

$$SA(F) = \sigma(f^{7\times 7}([AvgPool(F), MaxPool(F)]))$$



GCSA-BiFPN

Based on the BiFPN module, GC module, and SA module, the GCSA-BiFPN is built. GCSA-BiFPN pays attention to the discriminative context and spatial information, greatly improving the performance.

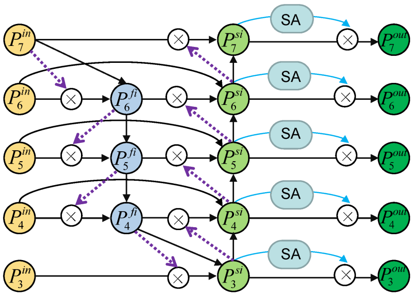
As a specific example, here we describe the fused features at level 6 for GCSA-BiFPN:

$$P_6^{fi} = Conv\left(\frac{GC(P_7^{in}) \otimes P_6^{in} + Resize(P_7^{in})}{2}\right)$$

$$P_6^{si} = Conv\left(\frac{P_6^{in} + GC(P_5^{si}) \otimes P_6^{fi} + Resize(P_5^{si})}{3}\right)$$

$$P_6^{out} = SA(P_6^{si}) \otimes P_6^{si}$$

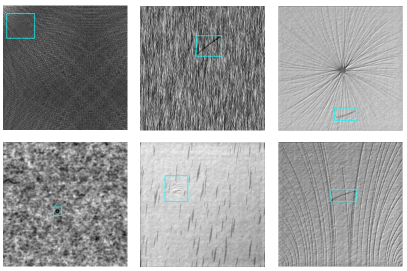
All the feature maps of other levels are constructed in a similar manner.



Results on DAGM2007 Dataset

The dataset consists of 10 classes of the defect in total, including 2100 defect images. Although the defect data is artificially generated, it is similar to the problem in the real world. Many types of defects are very similar to the common defect types of mobile phones.

Methods	Backbone	AP_{50}	AP_{75}
Two-stage:			
Faster R-CNN	ResNet-50	92.71	26.22
Faster R-CNN	ResNet-101	92.89	43.08
FPN	ResNet-50	96.99	62.85
FPN	ResNet-101	96.85	62.03
Cascade R-CNN w/FPN	ResNet-50	96.96	71.91
Cascade R-CNN w/FPN	**ResNet-101**	**97.44**	**74.45**
One-stage:			
Yolo v3	Darknet-53	87.81	25.97
RetinaNet	ResNet-50	97.10	61.40
RetinaNet	ResNet-101	95.40	64.50
EDD-Net D0(ours)	EfficientNet-B0	95.40	56.40
EDD-Net D1(ours)	**EfficientNet-B1**	**97.10**	**71.20**
EDD-Net D2(ours)	EfficientNet-B2	96.00	66.50



Results on MPSOSD Dataset

We group models together if they have similar accuracy on the MPSOSD dataset, and compare the performance between our EDD-Net and other detectors in each group. #Params and #FLOPs denote the number of parameters and the number of multiply-adds. Notably, our EDD-Net achieves better accuracy and efficiency than previous detectors on the MPSOSD dataset.

Methods	Backbone	AP_{50}	AP_{75}	#Params	#FLOPs
Yolo v3	Darknet-53	85.14	42.80	61.52M	32.76B
Cascade R-CNN w/FPN	ResNet-50	84.93	83.92	69.70M	707.66B
Cascade R-CNN w/FPN	ResNet-101	85.83	78.29	88.64M	897.06B
EfficientDet-D0	EfficientNet-B0	90.20	**87.70**	**3.83M**	**2.29B**
RetinaNet	ResNet-50	**93.30**	82.90	36.33M	74.17B
EDD-Net D0(ours)	**EfficientNet-B0**	**94.10**	**85.20**	**3.83M**	**2.30B**
RetinaNet	ResNet-101	94.70	92.30	55.32M	101.75B
EfficientDet-D1	EfficientNet-B1	95.30	**94.10**	**6.56M**	**5.58B**
Faster R-CNN	ResNet-50	95.63	55.02	27.99M	155.39B
FPN	ResNet-101	95.69	63.56	60.24M	432.37B
FPN	ResNet-50	96.24	60.74	41.30M	299.79B
Faster R-CNN	ResNet-101	**97.36**	61.64	46.94M	250.04B
EDD-Net D1(ours)	**EfficientNet-B1**	**98.40**	**94.70**	**6.56M**	**5.59B**
EfficientDet-D2	EfficientNet-B2	**98.40**	**97.60**	**8.01M**	**10.02B**
EDD-Net D2(ours)	**EfficientNet-B2**	**99.50**	**91.90**	**8.01M**	**10.03B**

Conclusion

To solve the small-scale and low contrast problem → GCSA-BiFPN

To achieve both high accuracy and better efficiency → EDD-Net D0-D2

To provide research worthy images for the field → MPSOSD Dataset

Acknowledgements

This work is supported by National Key R&D Program of China (2018YFB1308600, 2018YFB1308602).

Contact: levigty@stu.pku.edu.cn