EDD-Net: An Efficient Defect Detection Network

Tianyu Guo
levigty@stu.pku.edu.cn

Linlin Zhang
catherinezll@pku.edu.cn

Runwei Ding
dingrunwei@pku.edu.cn

Ge Yang
yangge@pkusz.edu.cn

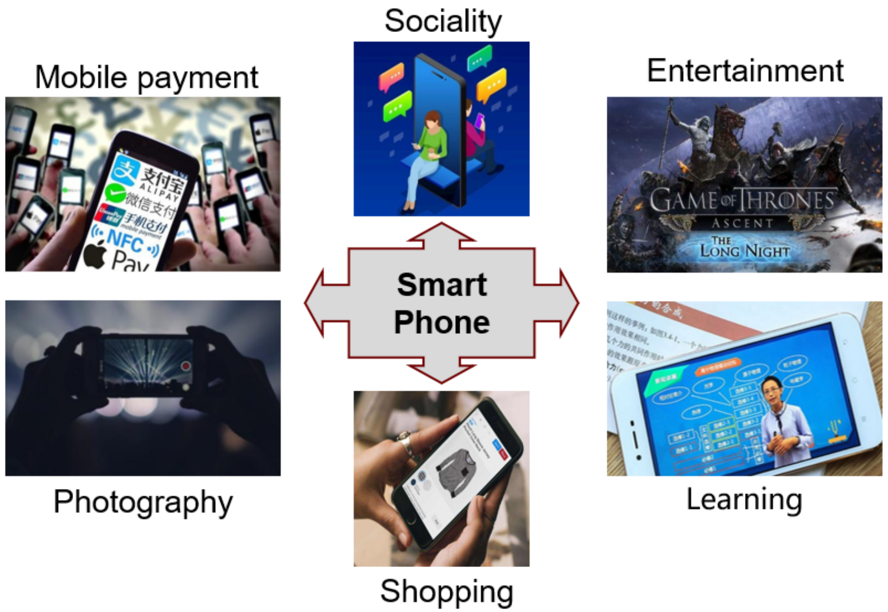
Key Laboratory of Machine Perception, Shenzhen Graduate School, Peking University

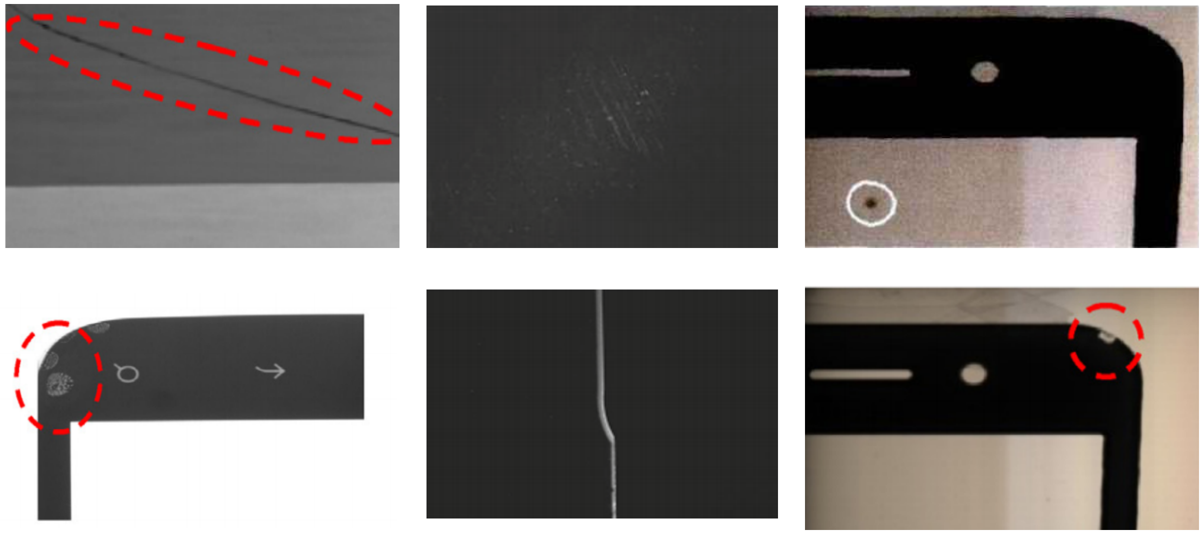
Presenter: Tianyu Guo

Introduction

PEKING UNIVERSITY

➢ Background

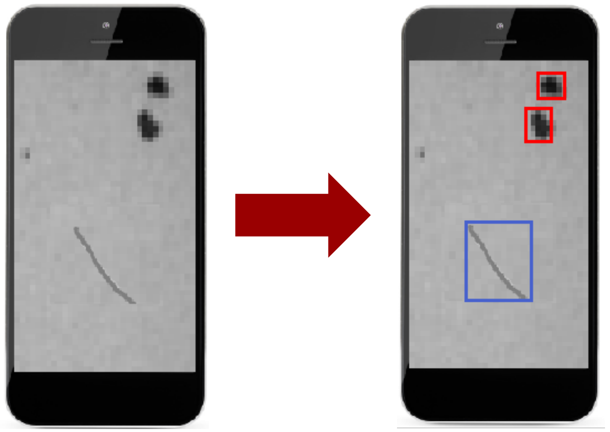




- Mobile phone has become an indispensable part of our daily life.
- The surface of the mobile phone directly affects the user experience.
- Various types of mobile phone surface defects inevitably exist on the production line.

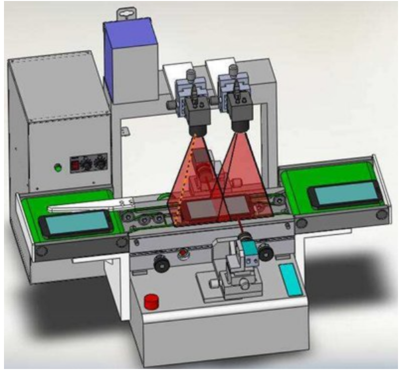
Introduction

➢ Task definition

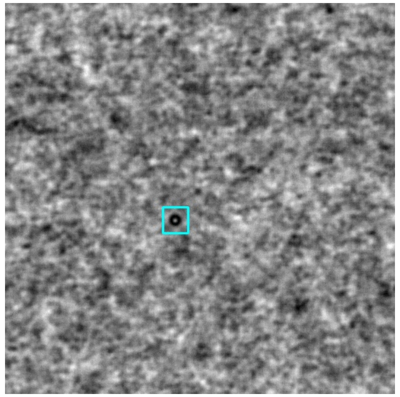


- Input: Mobile phone surface image
- Output: Defect location and defect type

➢ Motivation



- The defect detection task is based on actual application scenarios and has high requirements for accuracy and real-time performance

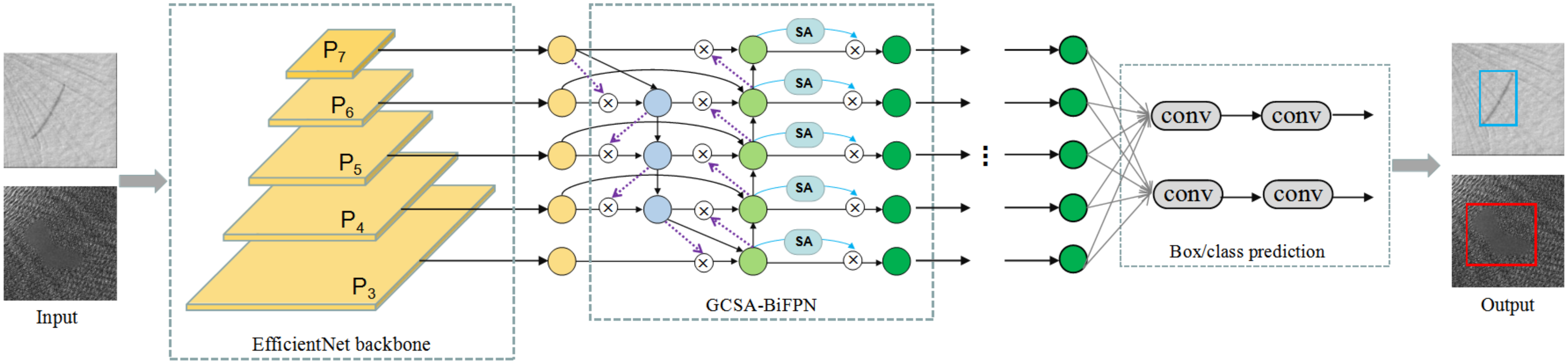


- Compared with object detection task, defect detection task has small-scale and low contrast challenge

We aim to propose efficient and accurate method to detect the defect target.

Method

➢ Pipeline of our proposed EDD-Net

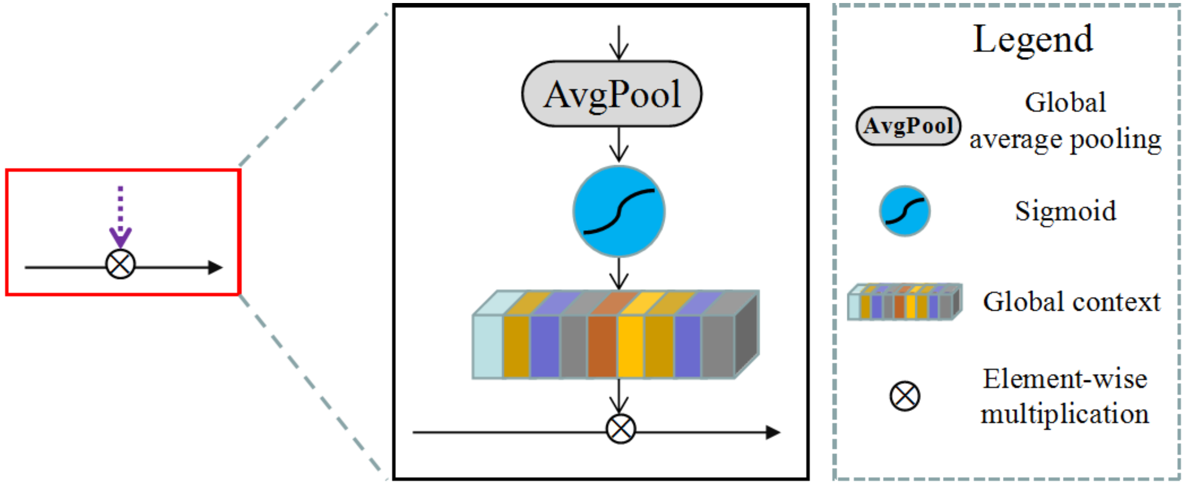


- EfficientNet ⟶ Extract features
- GCSA-BiFPN ⟶ Obtain more discriminative features
- Box/Class Prediction Network ⟶ Obtain the final results

	Input size	Backbone	GCSA-BiFPN channels	GCSA-BiFPN layers	Box/class layers
EDD-Net D0	512 × 512	EfficientNet-B0	64	3	3
EDD-Net D1	640 × 640	EfficientNet-B1	88	4	3
EDD-Net D2	768 × 768	EfficientNet-B2	112	5	3

Method

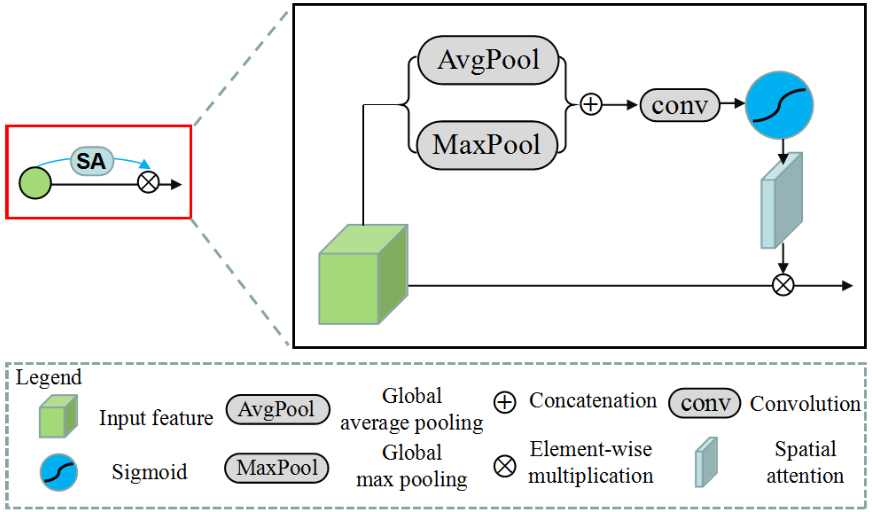
➢ Global Context Module



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

The GC module use global average pooling to generate global context information, and generate weighted vectors after the sigmoid function to guide the feature generation of the next level.

➢ Spatial Attention Module

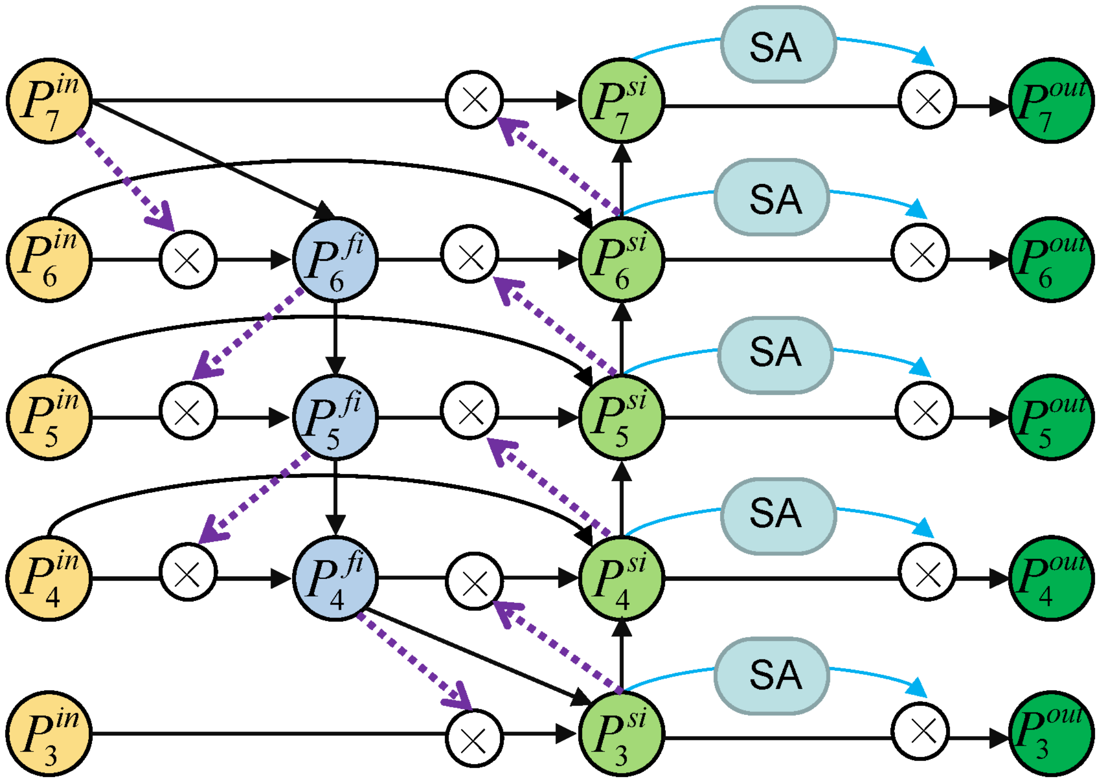


$$SA(F) = \sigma(f^{7\times 7}([AvgPool(F), MaxPool(F)]))$$

The Spatial Attention module first combines the spatial information generated by global average pooling and global max pooling. Then the spatial attention is generated after convolution and sigmoid function to optimize the original feature representation.

Method

➢ GCSA-BiFPN



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

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

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

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

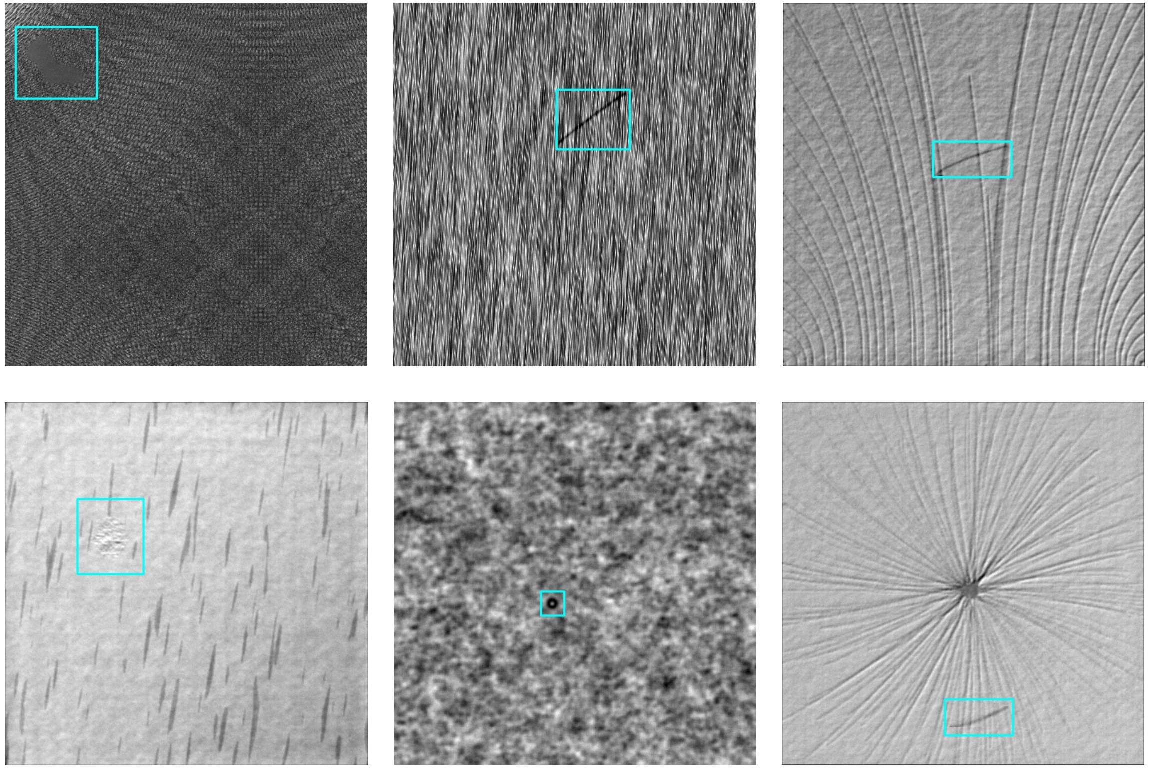
All the feature maps of other levels are constructed in a similar manner.

Experimental results

北京大学 PEKING UNIVERSITY

➢ DAGM2007 Dataset

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

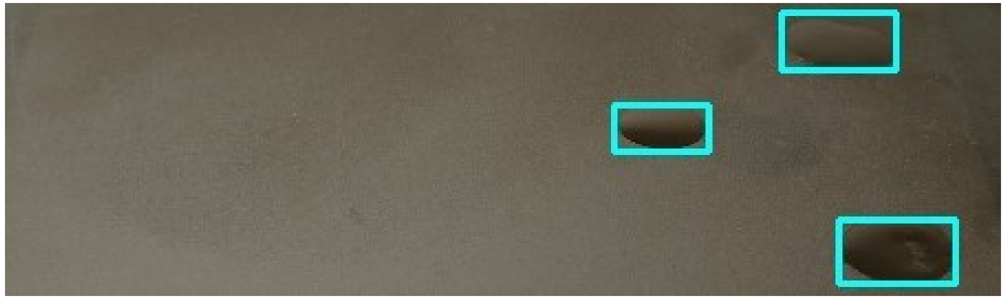


- The dataset consists of 10 classes of the defect.
- Our EDD-Net performs quit well on this dataset.

Experimental results

➢ 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**



- Our EDD-Net D0-D2 has well performance in both accuracy and efficiency

Conclusion

Focusing on the difference between defect detection task and the object detection task, EDD-Net has three novel aspects to adapt to defect detection.

- Firstly, EDD-Net is under the framework of advanced object detector EfficientDet which guarantees the detection results.
- Secondly, a novel feature pyramid module GCSA-BiFPN is proposed to fully use the context information and spatial information.
- Thirdly, our EDD-Net D0-D2 can easily achieve real-time performance while the accuracy is considerable, which has practical significance in different scenarios on the production line.

THANKS

Presenter: Tianyu Guo