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.

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:
\[ S(A(F)) = \sigma(F^T \sigma([AvgPool(F), MaxPool(F)])) \]

Spatial Attention Module

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:
\[ P^6 = \text{Conv}(GC(P^6) \odot P^6 + \text{Resize}(P^6)) \]
\[ P_{\alpha}^6 = \text{Conv}(P_{\alpha}^6 + GC(P^6) \odot P^6 + \text{Resize}(P^6)) \]
\[ P_{\alpha}^6 = SA(P_{\alpha}^6) \odot P_{\alpha}^6 \]

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.

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.

Conclusion
To solve the small-scale and low contrast problem

To achieve both high accuracy and better efficiency

To provide research worthy images for the field

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

Contact: levigty@stu.pku.edu.cn

 MPSOSD Dataset

GCSA-BiFPN

EDD-Net D0-D2

EDD-Net D1 (ours)
EfficientNet-B1
98.40 97.60 8.01M 10.03B

EDD-Net D2 (ours)
EfficientNet-B2
99.50 91.90 8.01M 10.03B

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.

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:
\[ S(A(F)) = \sigma(F^T \sigma([AvgPool(F), MaxPool(F)])) \]

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.

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.

Conclusion
To solve the small-scale and low contrast problem

To achieve both high accuracy and better efficiency

To provide research worthy images for the field

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

Contact: levigty@stu.pku.edu.cn

 MPSOSD Dataset

GCSA-BiFPN

EDD-Net D0-D2

EDD-Net D1 (ours)
EfficientNet-B1
98.40 97.60 8.01M 10.03B

EDD-Net D2 (ours)
EfficientNet-B2
99.50 91.90 8.01M 10.03B

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.

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:
\[ S(A(F)) = \sigma(F^T \sigma([AvgPool(F), MaxPool(F)])) \]

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.

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.

Conclusion
To solve the small-scale and low contrast problem

To achieve both high accuracy and better efficiency

To provide research worthy images for the field

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

Contact: levigty@stu.pku.edu.cn

 MPSOSD Dataset

GCSA-BiFPN

EDD-Net D0-D2

EDD-Net D1 (ours)
EfficientNet-B1
98.40 97.60 8.01M 10.03B

EDD-Net D2 (ours)
EfficientNet-B2
99.50 91.90 8.01M 10.03B

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.

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:
\[ S(A(F)) = \sigma(F^T \sigma([AvgPool(F), MaxPool(F)])) \]

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.

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.

Conclusion
To solve the small-scale and low contrast problem

To achieve both high accuracy and better efficiency

To provide research worthy images for the field

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

Contact: levigty@stu.pku.edu.cn

 MPSOSD Dataset

GCSA-BiFPN

EDD-Net D0-D2

EDD-Net D1 (ours)
EfficientNet-B1
98.40 97.60 8.01M 10.03B

EDD-Net D2 (ours)
EfficientNet-B2
99.50 91.90 8.01M 10.03B

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.

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:
\[ S(A(F)) = \sigma(F^T \sigma([AvgPool(F), MaxPool(F)])) \]