Screw hole location for laptop based on improved RCF

Yafei Huang, Jiwei Hu and Ping Lou

1 Department of Information and Communication Engineering, Wuhan University of Technology, Wuhan, Hubei, 430070, China
2 Department of Information and Communication Engineering, Wuhan University of Technology, Wuhan, Hubei, 430070, China
*hujjwei@whut.edu.cn

Abstract. The recent application of computer vision to disassembly and remanufacturing has pushed forward this field significantly. In disassembling the laptop, we find that the existing positioning screw methods such as hough transform and template matching were not suitable for black laptops, so we propose a new method to solve this problem. We use RCF network to locate the screw hole contour, and make two improvements according to the characteristics of small screw hole size: 1. A feature integration module is added to RCF network to increase multi-scale features. 2. A modulating factor is added to the loss function to enhance the attention to the indistinguishable pixels. Our approach has shown good results on black laptops and has higher ODS and OIS than the original RCF network and the existing method.

1. Introduction

With the improvement of production and living standards, notebook computers have become one of the necessities of daily life. The recycling of notebook computer parts is of great significance in the remanufacturing of waste electronics. Robotics could revolutionize manufacturing. Since most industrial robots cannot work in an unstructured environment, the costs can be greatly reduced if a robot can understand the tasks it needs to perform in a non-engineering environment through computer vision.

Due to the wide variety of notebook computers on the market and the fact that old computer screws can wear out, traditional machine vision methods are more difficult to cope with changeable situations. At the same time, the size and number of screws in the images taken by the robot camera are small, which will bring difficulties to the positioning of computer vision. Therefore, this paper proposes an improved RCF edge segmentation network to solve the problem of screw hole positioning. We add feature integration module to RCF edge segmentation network to add multi-scale features. The module is composed of global attention and local restructuring, across different spatial location and scale, collects task oriented features in the global and local, and improves test result of screw holes. In order to increase the attention to such indistinguishable edge features, we add the modulating factor into the loss function to control the shared weight of positive and negative samples to the total loss, so that the model could focus more on the difficult samples in training.

To sum up, our contributions are in three aspects:
1. A feature integration module is added to the RCF network to increase attention to small targets and realize the contour location of screw holes.
2. The modulating factor is added to the loss function to reduce the loss weight of the easily classified part to enhance the learning ability of the weak edge.
3. Experimental results show that our method is superior to the present one.
2. Related work
Screw positioning Nicholas M. DiFilippo and Musa K. Jouaneh [1] used Prewitt edge detection to enlarge the screw hole edges of laptop image. [2] used template matching and Generalized Hough Transform to recognize the bolts in the desired circular array. [3] used Douglas-Peucker algorithm and template matching to perform polygon approximation for contour segments. [4] through the application of Haar cascade, the problem of Phillips head screw detection based on camera was solved. [5] adopt the template matching method to detect the screws on the battery of electric vehicles.

Edge detection Contour data was grouped into child-classes, which learned the model parameters to fit each child-class [6]. A competent and precise edge detector, HED, was developed currently [7], which can carry out image-to-image training and prediction. All side output layers of the monolithic nested architecture were connected to the last convolution layer of each stage in VGG16 [8]. In addition, the relaxed labels generated from bottom-up edges were used to control the training process of HED [9]. The path was refined from top to bottom to learn clear edges effectively [10]. A hierarchical deep model learned the fusion of edge representations at different scales stably [11]. Richer convolution feature (RCF) were proposed [12], which performs edge detection of pixel prediction in image-to-image mode.

Small target detection Using the pyramidal feature hierarchy of convolutional network, feature pyramids are always constructed from low to high semantic level[13].Bharat Singh and Mahyar Najibi [14] determined Valid GT and invalid GT as well as valid anchor and invalid anchor in RPN network, so the pre-selected boxes can be more accurate. Wei Liu, Shengcai Liao [15] trained positioning modules by increasing the IOU threshold, so as to improve the positioning accuracy of small targets. Jianan Li, Xiaodan Liang[16] used Perceptual GAN to improve the feature representation of smaller objects, and made them similar to large objects.

3. Method
We collect 300 pictures of screws on the back panel of different laptops and expanded the data. The public edge data set is used for pre-training on the improved RCF network, and then the self-made screw data set is used to fine-tune the pre-training model based on transfer learning, so as to obtain the edge contour of the screw hole. Morphological processing is carried out on the contour map to solve the situation that there may be edge burrs and hole inside. The least square method is used to fit the circular contour of the screw hole, and the coordinates of the center point and radius of the hole were obtained. The following sections are described in detail.

3.1. Improved RCF network architecture
In RCF, all convolution features are well-package into a final representation in a holistic way that can be trained through back propagation. In this paper, we reintegrate the features of the final representation of RCF, formulating the feature integration module as global attention and local reconfiguration problems. The network structure will be show in figure 1. X1 ~ X5 is the part of five stage conv layers of RCF before up-sampling. We implement global feature module and local feature module by a light-weight network, so they could be embedded into the RCF network and learned end-to-end. The global and local operations complement to each other mutually because they deal with different scales of character hierarchies.
We use global section to emphasize information features and to suppress less useful features globally within a specific scope. The Squeeze-and-Excitation block [17] was applied as the base module for the global part, with more emphasis on the characteristics of the channel hierarchy. For the $x_1$ layer, every channel of the final representation on RCF operates a global pooling in squeeze module. The excitation stage is made up of two fully-connected layers which is followed by sigmoid activation.

The $\otimes$ denotes channel-wise multiplication. SE block takes into account the hierarchical information of the input image, which contributes to the ability to recognize features and more useful information globally. The feature level patch was mapped to an output feature patch, and shared with all local receptive fields in the local reconfiguration network. The local part slides the operation over the input to obtain the output feature maps. Here $\oplus$ denotes element-wise summation. This module enriches the hierarchical features and is beneficial to the location of screw holes.

### 3.2. Loss Function

In each image, all the ground truth was averaged to generate an edge probability graph from 0 to 1. In this case, 0 means there is no marker at this pixel, and 1 means there has a marker at this pixel. The pixels we marked are considered as positive samples while other pixels are considered as negative samples. In our screw hole data set, negative samples which are classified easily account for the most of the loss and dominate the gradient. While $\alpha$ and $\beta$ balance the influence of positive and negative samples in RCF, they do not distinguish between easy and hard samples. Instead, we recommend reshaping the loss function to reduce the weight of the easy samples, thus focusing the training on hard negatives. We propose to add a modulation factor $(1 - p_i)\gamma$ in the cross entropy loss. The parameter $\gamma$ is adjustable but no less than zero. The loss of each pixel we define relative to its label as:

$$I(X_i; W) = \begin{cases} \alpha \cdot (1 - P(X_i; W))^\gamma \log (1 - P(X_i; W)) & s_i = 0 \\ \beta \cdot (1 - P(X_i; W))^{\gamma} \log P(X_i; W) & s_i = 1 \end{cases}$$

in the equation

$$\alpha = \lambda \frac{|S'|}{|S'| + |S|}, \quad \beta = \frac{|S|}{|S'| + |S|}$$

$S'$ and $S$ are used as the total number of positive samples and the total number of negative samples respectively, which are balanced by the hyper-parameter $\lambda$. $X_i$ is the ground truth edge probability at pixel $i$. $P(X)$ represents the class of the final prediction in the model, and $W$ is the parameters learned in
RCF network. Therefore, our improved loss function is represented as equation 3, in which $X_i^{\text{fuse}}$ comes from the $1 \times 1$ conv layer after concat. $|I|$ is the number of pixels in image I. Thus, we can not only solve the problems caused by unbalanced samples, but also increase the attention to difficult samples such as screw hole edge.

$$L(W) = \sum_{i=1}^{\mid I \mid} (l(X_i^{\text{fuse}}; W))$$

(3)

3.3. Contour positioning based on transfer learning
We address the data set shortage by fine-tuning the improved RCF network. The public edge data set was used as the source domain to train on the improved RCF location network. After training, we use the expanded laptop screw hole data set as the target field, fine-tuning the training model. There are five sets of convolution layers in the improved network. In order to avoid overfitting caused by small amount of dataset in the target domain, the weights of conv layers in first two stages are initialized from the pre-trained improved model and froze while fine-tuning.

3.4. Morphological processing for coordinates
Since the contour of the screw hole found in the experiment may have edge burring and hole inside, we used such morphological processing methods as erosion and dilation to solve the problem. For the processed hole contour, the least square method is used to fit the standard circle, from which the coordinates of the center point and the radius of the screw hole can be obtained.

4. Experiment
We carry out seven groups of experiments. The loss function used in the fourth and fifth groups is RCF loss function, and our loss function is used in the last two groups, where $\gamma$ is 2. The experimental results will be shown in the following table 1. From table 1, we can see that adding feature integration modules to RCF alone and using our improved loss function on RCF alone can both achieve better results than the original RCF on the screw hole dataset. Improving together works best. From figure 2 we can find that our method of improving the network is slightly better than that of improving the loss function.

| method          | ODS  | OIS  |
|-----------------|------|------|
| HED[7]          | 0.709| 0.722|
| RDS[9]          | 0.715| 0.726|
| CED[10]         | 0.718| 0.731|
| RCF[12]         | 0.731| 0.743|
| Improved RCF    | 0.747| 0.760|
| RCF + improved loss | 0.745| 0.757|
| Improved RCF + improved loss | **0.758**| **0.771**|

Table 2. Varying $\gamma$ for loss.

| $\gamma$ | ODS  | OIS  |
|----------|------|------|
| 0        | 0.731| 0.743|
| 0.2      | 0.733| 0.745|
| 0.5      | 0.735| 0.748|
| 1.0      | 0.739| 0.750|
| 2.0      | **0.745**| **0.757**|
| 5.0      | 0.734| 0.746|

In the table 2, we find that the ODS and OIS get the best effect on the screw hole dataset when $\gamma$ is 2. $\gamma = 0$ has the same effect as the result training on original RCF network. Among the figure 3, the first column is the result trained by RCF, the second is our result, and the last is our edge map.

Figure 2. The results of screw hole dataset by different method.
5. Conclusion
In the experiment of screw hole positioning, we found that the small size of screw hole has an influence on the positioning effect. In this paper, we propose two methods to improve and solve it. We add a feature integration module in the RCF network to add multi-scale features, and add a modulating factor in the loss function to increase the attention to indistinguishable pixels. We have proved the validity of the two methods through experiments.

Acknowledgments
This project is supported by National Natural Science Foundation of China (Grant No. 52075404) and Application foundation frontier special project of Wuhan Science and Technology Bureau (No.2020010601012176).

Reference
[1] DiFilippo, N. M., Jouaneh, M. K. (2018). A System Combining Force and Vision Sensing for Automated Screw Removal on Laptops. IEEE Transactions on Automation Science and Engineering, 15(2), 887-895.
[2] U. Büker, S. Drüe, N. Götze, et al. Vision-based control of an autonomous disassembly station. 2001, 35(3-4):179-189.
[3] P. Gil, J. Pomares, S.v.T. Puente C. Diaz, et al. Flexible multi-sensorial system for automatic disassembly using cooperative robots. 2007, 20(8):757-772.
[4] Supachai Vongbunyong, Sami Kara, Maurice Pagnucco. Application of cognitive robotics in disassembly of products. 2013, 62(1):31-34.
[5] Gerbers, R., Wegener, K., Dietrich, F., & Dröder, K. (2018). Safe, Flexible and Productive Human-Robot-Collaboration for Disassembly of Lithium-Ion Batteries. In Recycling of Lithium-Ion Batteries 2018.pp. 99-126.
[6] W. Shen, X. Wang, Y. Wang, X. Bai, and Z. Zhang, 2015.Deep Contour: A deep convolutional feature learned by positive-sharing loss for contour detection, In: Computer Vision and Pattern Recognition.pp .3982–3991.
[7] Saining Xie, Zhuowen Tu,2017. Holistically-Nested Edge Detection. In: Computer Vision and Pattern Recognition.125(1-3):3-18.
[8] K.Simonyan and A.Zisserman,2015.Very deep convolutional networks for large scale image recognition. In Proc.Int.Conf.Learn.Representations.
[9] Y. Liu and M. S. Lew, 2016.Learning relaxed deep supervision for better edge detection, In Proc.IEEE Conf. Comput.Vis.Pattern Recognit 2016,pp.231–240.
[10] Y. Wang, X. Zhao, Y. Li, and K. Huang, “Deep crisp boundaries: From boundaries to higher-
level tasks,” IEEE Trans. Image Process., vol. 28, no. 3, pp. 1285–1298, 2019.

[11] D. Xu, W. Ouyang, X. Alameda-Pineda, E. Ricci, X. Wang, and N. Sebe, “Learning deep structured multi-scale features using attention-gated CRFs for contour prediction,” in Proc. Int. Conf. Neural Inf. Process. Syst., 2017, pp. 3961–3970.

[12] Liu Yun, Cheng Ming-Ming, Hu Xiaowei, et al. Richer Convolutional Features for Edge Detection. 2019, 41(8):1939-1946.

[13] Sun, F., Kong, T., Huang, W., Tan, C., Fang, B., & Liu, H. (2019). Feature Pyramid Reconfiguration with Consistent Loss for Object Detection. IEEE Trans Image Process.

[14] Bharat Singh, Mahyar Najibi, Larry S, 2018. Davis. SNIPER: Efficient Multi-Scale Training. In: Computer Vision and Pattern Recognition. 2018.

[15] Liu Wei, Liao Shengcai, Hu Weidong, Efficient Single-Stage Pedestrian Detector by Asymptotic Localization Fitting and Multi-Scale Context Encoding. 2019.

[16] Li, J., Liang, X., Wei, Y., Xu, T., Feng, J., & Yan, S. 2017. Perceptual Generative Adversarial Networks for Small Object Detection. In: Computer Vision and Pattern Recognition 2017.

[17] Hu Jie, Shen Li, Albanie Samuel, et al. Squeeze-and-Excitation Networks. 2020, 42(8):2011-2023.