# Small Data-Driven Electrical Insulator Defect Detection

Yuxin Song\* and Dingkai Susun\* Lei Pan<sup> $\dagger$ </sup> § and Ming Wu\* and Shengli Zhu<sup> $\ddagger$ </sup> and Hui Ma<sup> $\ddagger$ </sup>

\* Beijing University of Posts and Telecommunications, Beijing, China

E-mail: songyuxin@bupt.edu.cn, susundingkai@bupt.edu.cn, wuming@bupt.edu.cn

<sup>†</sup> Institute of Microelectronics of the Chinese academy of Sciences, Beijing, China

E-mail: panlei@ime.ac.cn

<sup>‡</sup> Beijing Ikingtec intelligent technology Co., Ltd, Beijing, China

E-mail: zsl@ikingtec.com, mh@ikingtec.com

<sup>§</sup> University of the Chinese academy of Sciences, Beijing, China

E-mail: panlei@ime.ac.cn

Abstract—The inspection and maintenance of insulator equipment has always adopted the traditional manual detection. It is very significant to study the automatic Insulator defect detection by drone inspection. However, in practical industrial applications, the number of available defect insulator samples is limited. It is difficult to construct a sufficient and high-quality dataset to support the training of the object detection model. In this paper, we propose a detection framework which combines the superresolution reconstruction and the object detection model. In our model, we use the super-resolution reconstruction and traditional data augmentation to amplify the amount of data and avoid the overfitting caused by the small sample data. The model has excellent performance on the training set which only contains 80 images, and achieves 61% mAP. We also show that the superresolution reconstruction can rich image texture features and is more effective than some traditional data augmentation methods.

### I. INTRODUCTION

Insulator equipment is an important part of the transmission line in the power grid, and it also play a good insulation role between the conductor, the cross bar and the tower. However, the insulator is very prone to failures during the long-time high-load operation and exposure to the natural environment, such as self-explosion, damage, and dropped strings.

The traditional insulator defect detection usually adopts the manual tower-to-tower inspection. With the rapid development of computer vision, the use of image acquisition equipment as a medium and deep learning technology for the defect detection of insulator equipment will greatly improve the detection efficiency and reduce the investment cost.

In this paper, using the insulator images taken by drone to detect the defective areas where the insulator is "selfexposed". However, in practical industrial applications, the number of available defect samples is limited. This leads to a difference between the distribution of the positive and the negative samples, which also makes it difficult to develop and deploy the traditional data-driven object detection models. Thus, the Data Augmentation is used to extend the few-shot dataset in a more efficient and controlled way. In addition to the traditional data augmentation, we also use the superresolution reconstruction method. On the one hand, super-

resolution can effectively expand the few-shot dataset, on the other hand, defective insulator is usually a small target compared to aerial image taken by drone. Super-resolution reconstruction makes up for the lack of pixel values well at the insulator defects area.

Emperiments on electrical insulator defect detection in a newly created dataset demonstrate that the proposed approach is able to train on a small number of defected samples dataset and also be applied to actual industrial detection.

## II. RELATED WORK

## A. Defect Detection

In recent years, deep learning have achieved good performance in various visual tasks [1-2]. Deep neural networks have shown excellent ability to extract image features. These deep image features are definitely helpful for defect extraction. Therefore, people have proposed several attempts to detect defects using deep neural networks. Yuan ZC et al. [3] proposed an improved segmentation network for the defect detection of the glass cover of the mobile phone external screen, and discussed the performance of adversarial network. Mei et al. [4] proposed approach which reconstructing image patches with convolutional denoising autoencoder networks at different Gaussian pyramid levels and synthesizing detection results from these different resolution channels. The effect is very good on the images with strong repetitive textures background such as cloth silk fabrics. However, the effect on metal surfaces and processed parts surface data sets is average or even poor. Feng C et al. [5] proposed a deep active learning system based on ResNet to maximize the performance. Xian Tao et al. [6] designed a new Cascaded Automatic Encoder (CASAE) structure for segmenting and locating defect. The cascading network combines a semantic segmentation for pixel-wise prediction and a classifier based on compact convolutional neural network (CNN).

## B. Insulator Defect Detection

Liao Shenglong [7] firstly uses the local features of the insulator for rough segmentation based on the active contour

model segmentation, and then uses the spatial distribution features of the insulator to detect the insulator. [8] proposed a recognition method that combine the shape, color and texture features of the insulator, perceives features from parallel lines in different directions of the insulator image as candidate regions of the insulator, and then expands from the candidate region to adjacent regions, using color features to locate the insulator. Finally, the average distance between the insulator pieces is used to divide the area into blocks, and the broken insulator is found by analyzing the texture features between the blocks. With the rapid development of deep learning in achieving end-to-end object detection, it has become possible to apply deep learning to the detection of insulator images. Chen Qing et al. [9] based on the convolutional neural network, first extracted the features of the insulator using the trained model, and then input the features into the self-organizing feature mapping network (SOM) for saliency detection. Miao et al. [10] used the SSD model to identify ceramic insulators and composite insulators from aerial images. Tao et al. [11] proposed a deep convolutional neural network (CNN) cascade architecture based on the regional proposal network(RPN), which successfully detect the insulator defects.

## C. Super-Resolution Reconstruction

The concept of SR first appeared in the field of optics. In the field of optics, SR refers to the process of trying to recover data outside the diffraction limit. Super-resolution reconstruction[12-13] uses the low resolution image sequence of the same scene to generate a high resolution image to effectively overcome the deficiencies of the hardware. Traditional super resolution reconstruction algorithm mainly include traditional interpolation amplification algorithms and learningbased super resolution reconstruction algorithms. The algorithm structure of traditional interpolation and amplification algorithms such as bicubic interpolation[14], lanczos interpolation, non-uniform interpolation[15-16], etc., is simple and easy to implement, but the reconstruction effect is poor. The learning-based super-resolution reconstruction usually have higher accuracy such as neighborhood embedding[17], sparse representation[18], regression learning [19], etc., but require a lot of training resources.

The SR algorithm based on convolutional neural network (CNN) has been widely used. Dong et al. [20,21] train a three layers deep convolutional network and use bicubic interpolation to reconstructed a low-resolution image. VDSR [22] proposed by Kim J et al. is the first method to introduce global residuals into SR. [23] proposed SRCNN which applied convolution layers on the pre-upscaled LR image. [24] proposed a super-resolution generative adversarial networks (SRGAN), which use deep residual network (ResNet) and skip connection. It use the series-parallel combined dilated convolution in the middle of the residual body to obtain different sizes scale information, and finally generate high-resolution reconstructed images.

#### III. PROPOSED METHOD

This paper proposes an insulator defect detection method for few-shot dataset. It is mainly divided into two parts: the first part uses super-resolution reconstruction to extends the our few-shot dataset, and the second part is defect detection using YOLO v3.

### A. Super-Resolution Reconstruction

SRGAN is a generative adversarial networks (GAN) for image super-resolution (SR) proposed by [24], which produces a reconstructions with high probability of realistic images. Our goal is to use a trained generator G to produce a highresolution image (SR) for a given low-resolution image (LR).

The generative model G is trained in order to fool the differentiable discriminator D which is trained to distinguish super-resolved images from real images. With the help of joint learning, the generator can learn to create solutions that are highly similar to real images and thus difficult to classify by D. Minimizing pixel-wise error measurements, such as the MSE, is chosen to compare the difference between the original images and the super-resolved images. Specifically, the generator network G uses residual blocks with identical layout in ResNet. The network combine two convolutional layers with small 3×3 kernels and 64 feature maps followed by batch normalization layers and ParametricReLU as the activation function. The SR images are produced by two trained sub-pixel convolution layers which act as a generator.

Discriminator network D is trained to discriminate real HR images from generated SR samples. It uses LeakyReLU activation and contains eight convolutional layers with an increasing number of  $3 \times 3$  filter kernels, increasing by a factor of 2 from 64 to 512 kernels as in the VGG network. Strided convolutions are used to reduce the image resolution each time the number of feature channels is doubled. The resulting 512 feature maps are followed by two dense layers and a final sigmoid activation function to obtain a probability for sample classification.

Our original data set contains 80 insulator images (OR) with resolution of 1024x768. First, resize the original dataset into a 256x192 low-resolution images (LR). Then use SRGAN to generate photo-realistic high-resolution images (SR) with high upscaling factors (4×). Finally, our new data set contains the original data set (OR) and the reconstructed image (SR) generated from the LR image. The example images that were interpolated and superresolved with a 4× upscaling factor are shown in Fig. 1.

SRGAN effectively alleviates the scarcity of defective samples, and can multiply enlarge our dataset through different high upscaling factors. This method fundamentally solve the contradiction between the small defect sample data and the large capacity of the object detection models. The generated SR images have high-quality at the defects. Under the premise that the shape and texture features of the original images have not changed significantly, appropriate noise is introduced into the images. This makes the trained detection model have strong robustness.



Fig. 1. Compared illustration of super-resolution and bicubic interpolation. (a) are the original images. (b) are the high-resolution images which resize from the low-resolution images (256x192). (c) are the super-resolution images which are reconstructed by SRGAN.

#### **B.** Defect Detection

In this paper, compared with the complex two-stage object detection model, the single-stage model often has a more lightweight network structure and fewer weights parameters. Therefore, in order to avoid overfitting caused by the small sample data during the training process, we chose YOLOv3[25] as the object detection model in the single-stage model. YOLOv3 [25] borrowed the idea of FPN [26] and select three features of different scales for merging, which makes the deeper convolutional layers are able to extract semantic features, but also retain the shallow features. Therefore, the small target still retains some high-quality surface features after passing through the continuously convolutional layers. It allows the defect of the insulator to be better detected by the network.

Similar to Section 3.A, in order to further solve the overfitting problem caused by a few-shot dataset, we use some data augmentation technologies such as horizontally flipped, translated, and randomly cropped on our synthesized dataset.

#### IV. EXPERIMENTS

### A. Few-Shot Dataset

We capture 80 images of defective insulators with resolution of 1024x768 by using a drone inspection to . Then according to 3.1, down-sampling each image with downscaling factors  $(4\times)$ , and using SRGAN to reconstruct LR images. The new SR images are mixed with the original aerial images as our new dataset. The final dataset contains three part. The first part is the original image, the second part is the HR (highresolution) image which is directly resized from LR (lowresolution) image by the BiCubic interpolation, the third part is SR (super-resolution) image which is constructed by SRGAN. Then some traditional data augmentation methods are used on the images in training set. Every image is horizontally flipped, randomly cropped and randomly translated. The dataset is further expanded after the traditional data augmentation.

#### B. Metric

In this paper, mAP (mean Average Precision) is used as the evaluation metric, that is, the average value of Precision under different Recall values is calculated according to the PR curve. Since the model only needs to detect one defect class, mAP is equivalent to AP, as in (1).

$$AP = \frac{1}{11} \sum_{\substack{r \in \{0, 0.1, \dots, 1.0\}\\\rho_{interp}(r) = \max_{\hat{r}: \hat{r} \ge r}} \rho_{interp}(r)$$
(1)

The accuracy rate represents the proportion of positive examples in the examples divided into positive examples, as in (2).

$$P = \frac{TP}{TP + FP} \tag{2}$$

Recall rate is a measure of coverage. Multiple positive examples of the metric are divided into positive examples, as in (3).

$$recall = \frac{TP}{TP + FN} \tag{3}$$

#### C. Experimental Details and Results

In the training phase, warmup is used to update the learning rate. The learning rate increases linearly from a small value until the specified epoch, and then decreases linearly. We set the initial learning rate from  $10^{-4}$  and the end learning rate from  $10^{-6}$ . The first training stage which increases the

learning rate continues 50 epochs and the second training stage continues 100 epochs. We set training batchsize from 2 because the image resolution ( $1024 \times 768$ ) is large compare to the memory of out GPUs (Nvidia 2080ti). For a mini batchsize, Warmup updating strategy reduces the model overfitting of mini-batch images in the initial stage of training, and keeps a stable distribution through all convolutional layers.

In order to prove the effectiveness of our data augmentation method, we designed three sets of experiments. The results shown in Table I. The training set AUG-1 contains 80 original images, and traditional data augmentation method (horizon-tally flipped, randomly cropped and randomly translated) are used on every image. The trained model is overfitting and the AP on test set is only 0.29. The training set NOAUG-3 contains original images, HR images which are directly resized from low-resolution images by the BiCubic interpolation and the SR images which are reconstructed from low-resolution images by SRGAN. The trained model without any traditional data augmentation has 0.52 AP. The final training set AUG-3 performs traditional data augmentation on each image based on NOAUG-3. The AP on AUG-3 is 9% higher than NOAUG-3.

TABLE I COMPARISON OF DIFFERENT DATA AUGMENTATION STRATEGIES (AUG-3,NOAUG-3,AUG-1)

Training Set	AP(Average Precision)
AUG-1	0.29
NOAUG-3	0.52
AUG-3	0.61

The AP of the final model on the test set is 61.11%, and the test effect diagram is illustrated in Fig. 2.



Fig. 2. Defect detection results in the part of test set.

#### V. CONCLUSIONS

In this paper, considering the scarcity of defective insulator samples, SRGAN is used to extend the dataset and enrich the training samples. The serious overfitting problem caused by the few-shot dataset in the training process is effectively solved by combining HR and SR images. The experimental results on the insulator defect dataset show that the effectiveness of the combination of SRGAN and YOLOv3. The proposed methods also can be applied to actual industrial detection.

#### ACKNOWLEDGMENT

This work was supported in part by the National Key R&D Program of China under Grant 2019YFF0303300 and under Subject II No. 2019YFF0303302, in part by the National Natural Science Foundation of China (NSFC) under Grant U19B2036, in part by MoE-CMCC "Artifical Intelligence" Project No. MCM20190701

#### REFERENCES

- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany. 2015. 234–241.
- [2] Kampffmeyer M, Salberg AB, Jenssen R. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops. LasVegas, NV, USA. 2016. 680–688.
- [3] Yuan ZC, Zhang ZT, Su H, et al. Vision-based defect detection for mobile phone cover glass using deep neural networks. International Journal of Precision Engineering and Manufacturing, 2018, 19(6): 801–810. [doi: 10.1007/s12541-018-0096-x]
- [4] Mei S, Yang H, Yin Z. An Unsupervised-Learning -Based Approach for Automated Defect Inspection on Textured Surfaces[J]. IEEE Transactions on Instrumentation and Measurement, 2018:1266-1277.
- [5] Feng C , Liu M Y , Kao C C , et al. Deep Active Learning for Civil Infrastructure Defect Detection and Classification[C]// ASCE International Workshop on Computing in Civil Engineering 2017. 2017.
- [6] Tao X , Zhang D , Ma W , et al. Automatic Metallic Surface Defect Detection and Recognition with Convolutional Neural Networks[J]. Applied Sciences-Basel, 2018, 8(9).
- [7] Liao Shenglong. Research on Key Technologies of Component Detection in Aerial Transmission Line Images [Master Thesis]. Dalian Maritime University, 2017 (in Chinese).
- [8] Wanguo Wang, Yufu Wang, Jun Han, et al. Recognition and Drop-Off Detection of Insulator Based on Aerial Image. 2016 9th International Symposium on Computational Intelligence and Design (ISCID). IEEE, 2016: 162 167.
- [9] Chen Qing, Yan Bin, Ye Run, et al. Research on Convolutional Neural Network Detection and Self-Explosion Recognition of Aerial Insulators. Journal of Electronic Measurement and Instrument, 2017(6): 942 943 (in Chinese).
- [10] Airen Miao, Xinyu Liu, Jing Chen, et al. Insulator Detection in Aerial Images for Transmission Line Inspection Using Single Shot Multibox Detector. IEEE Access 2019,7: 9945 9956.
- [11] Xian Tao, Dapeng Zhang, Zihao Wang, et al. Detection of Power Line Insulator Defects Using Aerial Images Analyzed With Convolutional Neural Networks. IEEE Transactions on Systems, Man, and Cybernetics Systems, 2018: 1 13.
- [12] DENG CH ZH, TIAN W, WANG SH Q, et al. Super-resolution reconstruction of approximate sparsity regularized infrared images[J]. Opt. Precision Eng. 2014, 22(6): 1648-1654. (in Chinese).
- [13] PENG ZH M, JING L, HE Y M, et al. Super-resolution fusion of multifocus image based on multi-scale sparse dictionary[J]. Opt. Precision Eng., 2014, 22(1): 169-176. (in Chinese).
- [14] Hou H S, Andrews H C. Cubic spline for image interpolation and digital filtering[J]. IEEE Transaction on signal processing, 1978, 6(6): 508-517.
- [15] Park S C, Park M K, Kang M G, Super resolution image reconstruction: a technical review[J]. IEEE Signal Processing Magazine, 2003: 21-35.
- [16] Aftab H, Mansoor A B. Asim M. A New Single Image Inter-polation Technique For Super Resolution[c]. //EEE International Multitopic Conference. 2008: 592-596.

- [17] Chang H, Yeung D Y, Xiong Y. Super—resolution Through Neighbor Embedding[c]. //EEE Computer Society Conference On Computer Vision And Pattern Recognition, 2004: 275-282.
- [18] Yang J C, Wright J, Huang T, et al. Image Super-Resojution As Sparse Representation Of Raw Image Patches[C]. //EEE Computer Society Conference On Computer Vision And Pattern Recognition, 2008: 1-8.
- [19] He H. Ch S W. Signal Image Super—resolution using Gaussian process regression[c]. //IEEE Computer Society Conference On Computer Vision And Pattern Recognition, 201 1: 449-456.
- [20] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In European Conference on Computer Vision (ECCV), pages 184–199. Springer, 2014. 3, 6, 8.
- [21] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2):295–307, 2016. 3, 5, 7.
- [22] Kim, J., Lee, J K., Lee, K M.: Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016, pp. 1646–1654 (2016).
- [23] Dong, C., Loy, C C., He, K., et al.: Learning a Deep Convolutional Network for Image Super-Resolution. In: European Conference on Computer Vision. Springer, Cham (2014).
- [24] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image superresolution using a generative adversarial network[C] //Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4681-4690.
- [25] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.
- [26] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.