Few-shot learning approach for 3D defect detection in lithium battery

Ke Wu1,2, Jie Tan1,*, Jingwei Li1,2, Chengbao Liu1

1Institute of Automation, Chinese Academy of Sciences, Beijing, China

2School of artificial intelligence, University of Chinese Academy of Sciences, Beijing, China

*Corresponding author’s e-mail: jie.tan@tom.com

Abstract. Detecting the surface defects in a lithium battery with an aluminium/steel shell is a difficult task. The effect of reflectivity, the limitation of acquiring the 3D information, and the shortage of massive amounts of labelled training data make the 2D detection method hard to classify surface defects. In this work, a few-shot learning approach for 3D defect detection in lithium batteries is proposed. The multi-exposure-based structured light method is introduced to reconstruct the 3D shape of the lithium battery. Then, the anomaly part of the 3D point cloud is transferred into 2D images by the height-gray transformation. The MiniImageNet datasets are used as the source domain to pretrain the Cross-Domain Few-Shot Learning (CD-FSL) model. The accuracy in our experiment result is 97.17%, which means that our method can be used to classify the surface defects of the lithium battery.

1. Introduction

As an important secondary energy source, lithium-ion batteries are becoming the most widely used energy storage substance. Lithium-ion batteries combine several important benefits such as high energy density, exhibiting no memory effects, having low self-discharge rate, and a relatively cheap production cost. The lithium battery is lightweight and powerful, at the same time, they are prone to leaking and catching fire [1-2]. Therefore, the monitoring of the production process and early detection of electrode defects are especially important. In the present industrial production, the appearance of lithium battery defects has a great impact on the battery qualification rate.

The appearance detection of lithium batteries with aluminium/steel shell defects nowadays has two main difficulties. On the one hand, the effect of the reflective surface and the limitation of the 2D computer vision detection methods make the detection become difficult. On the other hand, the number of defective lithium battery samples is too small to train the deep neural network.

A big amount of experts have already conducted a series of studies on this issue. There is an increasingly clear trend for algorithms based on deep learning methods to substitute machine learning algorithms for the classification of lithium battery defects. Regrettably, the lack of large datasets with dependable annotations hinders the development of such algorithms, especially for the particular lithium battery defects. [2-4]. Recently, a subfield known as Few-Shot Learning (FSL) has been demonstrated the effectiveness of several architectures to learn new classes using few labelled data[5-8].

In many cases, it is very hard, costly, or even not possible to collect large amounts of labelled data, such as medical data, data manually labelled by users on cell phones. The same is true for product defects in the industry. The ability to train with a small amount of labelled data and then get a good model with
more satisfactory usage has become a very important theme in the development of machine learning and has received a lot of attention from academia and industry. Few-shot learning is used to identify the sameness and similarity between targets from few samples. (not required to have appeared in the training data set) [9-11]. For example, the neural networks will be able to learn new classes from the labelled datasets once they have learnt to compare classes [12]. Numbers of metric learning algorithms have been explored including centroids based method, dimensionality reduction based method and so on[13]. Euclidean distance to class-mean representations [14], CNN relation modules [15], ridge regression [16], and graph neural networks [17]. In addition, the data shortage can be overcome by data augmentation [18], or use generative adversarial networks [19] or autoencoders [20].

In the actual field of industrial defect detection, there are a large number of situations where 3D information such as object height, surface angle, flatness, thickness, volume, etc. is required. Since 2D vision cannot obtain spatial coordinate information of objects, it does not support features such as measurements related to 3D shapes, or distinguishing objects of the same color. Moreover, 2D vision is particularly dependent on illumination and color/grayscale changes, and measurement accuracy is easily affected by illumination conditions, so it cannot be applied to defect detection and classification of high-gloss surfaces either.

To deal with the problem mentioned above. In this work, a few-shot learning approach for 3D defect detection in a lithium battery is proposed for the classification of the lithium battery surface defects. The exposure fusion-based phase shift method is adopted to accurately reconstruct the 3D shape of the lithium battery with a reflective surface. With the accurate 3D shape, the anomaly part of the defective lithium battery can be found by comparing it with the standard lithium battery model. The selected lithium battery can be classified well by designing the cross-domain few-shot learning model.

2. Approach

2.1. Multi-exposure based phase shift method

2.1.1. Phase Shift method.

The phase shift method is utilized as a 3D reconstruction method in this work. Generally, image acquisition can be divided into 3 steps [21]: fringe projection, fringe reflection, and fringe acquisition. Mathematically, a typical fringe pattern of the phase shift method can be represented as:

$$I(x, y) = A(x, y) + B(x, y)\cos\left[\phi(x, y) + \delta_k\right]$$

(1)

Where $A(x, y)$ represents the intercept value, $B(x, y)$ represents the amplitude, $x$ and $y$ represents the pixel horizontal coordinate and vertical coordinate of the projector respectively, $\phi$ is the phase that we want to figure out, and $\delta_k$ represents the phase shift. Generally, we set $\delta_k = 2\pi (k - 1)/3$. Then the phase $\phi$ can be computed as:

$$\phi(x, y) = \arctan\left\{\frac{\sqrt{3} [I_1(x, y) - I_3(x, y)]}{2I_2(x, y) - I_1(x, y) - I_3(x, y)}\right\}$$

(2)

The value of $\phi(x, y)$, ranging from $[-\pi, \pi]$ with $2\pi$ discontinuities, which is usually called the wrapped phase. By utilizing the phase unwrapping method, continuous phase map can be acquired from the discontinuities as mentioned above [22]. If the continuous phase map is retrieved, the 3D reconstruction result of the target can be obtained from this phase after calibration.
2.1.2. Exposure fusion method
Exposure fusion is a method that fusing a series of images into a well exposure image. Generally, this method is used in the case that the target has a high dynamic range reflectivity surface when reconstructing its 3D shape. Three main quality measures: the contrast, the saturation, and the well-exposedness of the input images need to be calculated at first. Then all of the input images are captured by the camera. At the end, the fused result can be acquired by collapsing the input images using weighted blending. Please see [6] for detail.

2.1.3. Height-Gray transformation
The feature is an important judgment basis for defect classification. Selecting remarkable features not only can significantly reduce the computing complexity, but also can dramatically improve the accuracy as well as the efficiency of defect identification. This process converts the height information to grayscale information and saves it in the picture. 3D information is well preserved and becomes an important basis for defect classification.

![The Defective lithium battery](image1.png)
![Height-gray transformed Image](image2.png)

Figure 1. The height-gray transformed image

For the target lithium battery as shown in Figure 1. Defining the 3D reconstruction result generated by multiple exposure fusion methods as a 3D cloud point $P_0$. The prepared standard 3D cloud point is $P_0$. Comparing 3D cloud point $P$ with $P_0$, and extracting the part that exceeds the height threshold $a_0$ to obtain the disparity point cloud $P_d$. Transforming the height value into an 8-bit gray value and obtaining the gray image $I_x$. There are three main common damages and defects of the lithium batteries: the dents, the bumps, and the scratches to be detected in our task. The selecting of the threshold value $a_0$ is related to the process requirement, if there is no grayscale point cloud with height information exceeding the threshold value between $P_0$ and $P$, it means that the inspection object corresponding to the point cloud is a qualified product in appearance, otherwise it means that the object is a defective product.

2.2. Cross-Domain Few-shot learning Model for Lithium battery defect classification

2.2.1. Baseline CNN
Figure 2 shows the schematics of the proposed cross-domain few-shot learning model for lithium battery defect classification. In this model a ResNet-10 Convolutional Neural Network (CNN) is adopted to obtain the image features $f^B$, and build a softmax predictor $C_1$ in this feature space. The MiniImageNet datasets are used as the source domain to pretrain the cross-domain few-shot learning model. In this stage, the labelled data in the MiniImageNet datasets are applied to pre-train the Res-Net-10, as well as...
the advanced feature adopter \( F^E(x) = f^B \), and the classifier \( C \), by minimizing the cross-entropy loss. In a training batch \((X_B, Y_B) = \{(X_i, Y_i) | i = 1, 2, \cdots, n\}\), the loss function can be written as:

\[
L_{\text{cross}}(C \circ F^E, X_B, Y_B) = \frac{1}{n} \sum_{i=1}^{n} L_{\text{cross}}(C \circ F^E(X_i), Y_i)
\] (3)

Where \( L_{\text{cross}} \) represents the cross-entropy loss function in this equation. This processing also can be conducted in the training stage of the target domain. Due to the shortage of the labelled data in the target domain. Generally, we just need to conduct the fine-tuning processing to the network. To obtain the final test result on the query datasets, we just need to take the prediction result produced by the classifier \( C \).

![Figure 2. Schematics of the proposed few-shot learning architectures.](image)

2.2.2. The Regularization of Batch Spectrum

In this work, a batch spectrum regularization (BSR) mechanism is used as the penalizer. This penalizer is leveraged to inhibit the singular value of the advanced feature matrices in the training stage. This mechanism has the advantage of avoiding over-fitting when training the source datasets, as well as improving the capability to generalize to the target datasets.

In particular, for a training algorithm based on stochastic gradient descent, we train the network with the batch data. Supposing we have a given input training data \((X_B, Y_B)\), and we define the advanced feature matrix as \( A = [f_1^E, \cdots, f_n^E] \), where \( n \) represents the size of the batch data. In addition, the feature vector \( f_i^E \) acquired by Res-Net-10 represents the \( i \)-th input data in the batch. At the end, the penalizer can be shown as:

\[
L_{\text{tor}}(A) = \sum_{i=1}^{n} \sigma_i^2
\] (4)
Where $A$ represents the feature matrix of the batch data, and $\sigma_1, \sigma_2, \cdots, \sigma_b$ represents singular values of $A$. According to the definition of the cross-entropy loss function and the batch spectrum regularization mechanism. The training loss for each batch can be written:

$$L = L_{\text{cross}} \left( C \circ F^E; X_{\beta}, Y_{\beta} \right) + \lambda L_{\text{bsr}} \left( F^E \left( X_{\beta} \right) \right)$$  \hspace{1cm} (5)

### 2.2.3. Label Propagation

In many cases, it is very hard, costly, or even not possible to collect large amounts of labelled data. Inspired by an efficient label refinement process in [24], the label propagation (LP) is adopted to capture the underlying distribution of the entire data which can help us overcome the shortage of labelled data. The main idea of the label propagation is that the similar data should share the same label. We define the matrix $\hat{Y}^0$ as the matrix which predict the score of the training instances with a fine-tuned classifier. The $X_Q$ represents the training instances and the $C_t$ represents the fine-tuned classifier.

The k-NN graph is also built to evaluate the distance between samples, including the labelled samples and the unlabelled samples. Here the squared Euclidean distance is adopted to evaluate the distance between each pair images in our work. And the distance can be described as:

$$d(i, j) = \left\| F^E(x_i) - F^E(x_j) \right\|^2$$

The weight between each pair nodes $(i, j)$ can be calculated as follows:

$$W_{ij} = \begin{cases} 
\exp \left( -\frac{d(i, j)}{2\alpha^2} \right) & \text{i, j are the k-nearest neighbours} \\
0 & \text{otherwise}
\end{cases}$$  \hspace{1cm} (7)

Where $\alpha$ is a hyper parameters in general. In this equation it represents the radius of the RBF kernel.

The normalized Laplacian matrix $L$ can then be calculated as $L = Q^{-\frac{1}{2}} W Q^{-\frac{1}{2}}$, where $Q$ represents a diagonal matrix, and $Q_{ii} = \sum_j P_{ij} W_{ij}$. The LP is then employed to obtain the refreshed score matrix:

$$Y^* = \left( I - \gamma L \right)^{-1} \hat{Y}^0$$  \hspace{1cm} (8)

Where $I$ is an identity matrix, and $\gamma$ is a trade-off parameter whose value is set between 0 to 1. $\hat{y}_i = \arg\max_j Y^*_{ij}$ is used as the predicted class for the $i$-th image after the label propagation is employed.

### 2.2.4. Data Augmentation

To avoid over-fitting when training our model, the data augmentation (DA) strategy is also exploited to enrich the support datasets. There are several random operations, for example the image flipping, image rotation, random scaling, color jittering and random crop are applied to the labelled images to generate a few different images. The changed support set by DA can be fine-tuned. The same enhancement can be performed for the query set, and several variants of each image can be generated to share the same label. Thus, the prediction results for each image can be identified by taking an average of the predictions of all augmented variants of the same image.
3. Experiments

3.1. Setup

There are two kinds of datasets in our experiments, the source domain and the target domain. The source domain $D_s$ has 60,000 images obtained from the MiniImagenet datasets. The images in the source domain $D_s$ is utilized to pre-train the baseline feature extraction CNN $E^F$. The target domain $D_t$ has 60 images and 3 kinds of the defective lithium batteries images. The images in the target domain $D_t$ is used to fine-tune the few-shot learning model. For each class there are 10 images are randomly selected as the query set. There are 30 images are randomly sampled from the target domain which is used to train the classifier $C_2$ as shown in Fig. 2.

As we have mentioned above, the ResNet-10 is utilized in the model as a feature extractor, which is represented by $F^E$ in Fig. 2. The all-connected layer is adopted as the classifier in this model, which is represented by $C$ in Fig. 2. The value of the parameters $\lambda$ and $\beta$ are set as 1.0e-3, 1.0e-1 respectively.

As for the k-NN graph we mentioned in LP, the value of $k$ and $\alpha$ are set as 10, 2 respectively.

The stochastic gradient decent is utilized as the loss calculating method in our model, and the value of the momentum is set as 0.9. When we pre-train our model in source domain, the number of the epochs is set as 4.0e2. The learning rate $r_s=1.0e-3$, the weight decay $w_d=5.0e-4$. When we fine-tune our model in target domain, the number of the epochs is set as 1.0e2. The learning rate $r_t=1.0e-2$.

There are five kinds of augmentation operations are adopted in the DA stage in our experiments. The DA operation, parameters, and serial number are listed in Tab. I. The specific DA operations adopted in each datasets are also listed in Tab. II.

| Serial No. | Operation        | Parameters       |
|------------|------------------|------------------|
| I          | Random Scaling   | $84 \times 84$   |
| II         | Random Crop      | $84 \times 84$   |
| III        | Image Jitter     | Brightness: 0.4; Contrast: 0.4; Color: 0.4 |
| IV         | Image flipping   | The probability to flip: 50% |
| V          | Image Rotation   | 0-45°            |

Table 2. The modes of DA operation

| Datasets                                      | Augmentation       |
|-----------------------------------------------|---------------------|
| ISIC & Euro-SAT & Crop Diseases               | I+I,III,IV,V+I,V+I,III+I,IV |
| Chest-X                                       | I+I,III,IV+II,III+I,IV |

3.2. Results

The generated height-gray images from the multi-exposure based phase shift method as mentioned in section II part 1 are partly shown in Figure 3. Three main common damages and defects of the lithium batteries: dents, bumps, and scratches are separately shown in Figure 3(a)-(c). The average accuracy of our proposed method achieves 97.17%, which shows that our method can be used to classify the surface defects of the lithium battery.
Figure 3. Three main common damages and defects of the lithium batteries. (a) dents; (b) bumps; (c) scratches

4. Conclusions
In this paper, a few-shot learning approach based structured light method for the classification of the lithium battery 3D defects is proposed. This study shows that our method can overcome the effect of reflectivity, the limitation of acquiring the 3D information, and the shortage of large amounts of labelled training data. The target domain accuracy is 97.17% and this experiment result shows that our method can be used for the classification of lithium batteries with an aluminium/steel shell.

Acknowledgments
This work was supported by the National Natural Science Foundation of China under Grant U1801263 and U1701262 of the People's Republic of China. (Corresponding authors: Jie Tan and Chengbao Liu)

References
[1] Nitta, N., Wu, F., Lee, J. T., & Yushin, G. (2015). Li-ion battery materials: present and future. Materials today, 18(5), 252-264.
[2] Mohanty, D., Hockaday, E., Li, J., Hensley, D. K., Daniel, C., & Wood III, D. L. (2016). Effect of electrode manufacturing defects on electrochemical performance of lithium-ion batteries: Cognizance of the battery failure sources. Journal of Power Sources, 312, 70-79.
[3] Badmos, O., Kopp, A., Bernthaler, T., & Schneider, G. (2020). Image-based defect detection in lithium-ion battery electrode using convolutional neural networks. Journal of Intelligent Manufacturing, 31(4), 885-897.
[4] Ma, L., Xie, W., & Zhang, Y. (2019). Blister defect detection based on convolutional neural network for polymer lithium-ion battery. Applied Sciences, 9(6), 1085.
[5] Argüeso, D., Picon, A., Irusta, U., Medela, A., San-Emeterio, M. G., Bereciartua, A., & Alvarez-Gila, A. (2020). Few-Shot Learning approach for plant disease classification using images taken in the field. Computers and Electronics in Agriculture, 175, 105542.
[6] Wu, J., Zhao, Z., Sun, C., Yan, R., & Chen, X. (2020). Few-shot transfer learning for intelligent fault diagnosis of machine. Measurement, 166, 108202.
[7] Li, X., Wu, J., Sun, Z., Ma, Z., Cao, J., & Xue, J. H. (2020). BSNet: Bi-Similarity Network for Few-shot Fine-grained Image Classification. IEEE Transactions on Image Processing.
[8] Ji, Z., Zou, X., Huang, T., & Wu, S. (2020). Unsupervised Few-Shot Feature Learning via Self-Supervised Training. Frontiers in computational neuroscience, 14.
[9] Finn, C., Xu, K., & Levine, S. (2018). Probabilistic model-agnostic meta-learning. arXiv preprint arXiv:1806.02817.
[10] Nichol, A., Schulman, J. (2018). Reptile: a scalable meta-learning algorithm. arXiv preprint arXiv:1803.02999
[11] Rusu, A. A., Rao, D., Sygnowski, J., Vinyals, O., Pascanu, R., Osindero, S., & Hadsell, R. (2018). Meta-learning with latent embedding optimization. arXiv preprint arXiv:1807.05960.
[12] Koch, G., Zemel, R., & Salakhutdinov, R. (2015, July). Siamese neural networks for one-shot image recognition. In ICML deep learning workshop (Vol. 2).
[13] Vinyals, O., Blundell, C., Lillicrap, T., Kavukcuoglu, K., & Wierstra, D. (2016). Matching networks for one shot learning. arXiv preprint arXiv:1606.04080.
[14] Snell, J., Swersky, K., & Zemel, R. S. (2017). Prototypical networks for few-shot learning. arXiv preprint arXiv:1703.05175.
[15] Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P. H., & Hospedales, T. M. (2018). Learning to compare: Relation network for few-shot learning. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1199-1208).
[16] Bertinetto, L., Henriques, J. F., Torr, P. H., & Vedaldi, A. (2018). Meta-learning with differentiable closed-form solvers. arXiv preprint arXiv:1805.08136.
[17] Kim, J., Kim, T., Kim, S., & Yoo, C. D. (2019). Edge-labeling graph neural network for few-shot learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 11-20).
[18] Hariharan, B., & Girshick, R. (2017). Low-shot visual recognition by shrinking and hallucinating features. In Proceedings of the IEEE International Conference on Computer Vision (pp. 3018-3027).
[19] Zhang, R., Che, T., Ghahramani, Z., Bengio, Y., & Song, Y. (2018). MetaGAN: An Adversarial Approach to Few-Shot Learning. NeurIPS, 2, 8.
[20] Schonfeld, E., Ebrahimi, S., Sinha, S., Darrell, T., & Akata, Z. (2019). Generalized zero-and few-shot learning via aligned variational autoencoders. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8247-8255).
[21] Ekstrand, L., & Zhang, S. (2011). Autoexposure for three-dimensional shape measurement using a digital-light-processing projector. Optical Engineering, 50(12), 123603.
[22] Zhang, S. (2018). Absolute phase retrieval methods for digital fringe projection profilometry: A review. Optics and Lasers in Engineering, 107, 28-37.
[23] Wu, K., Tan, J., Xia, H., & Liu, C. (2020). An exposure fusion-based structured light approach for the 3D measurement of a specular surface. IEEE Sensors Journal.
[24] Ye, M., & Guo, Y. (2017, January). Labelless Scene Classification with Semantic Matching. In BMVC.