Triplet Siamese Network Model for Lithium-ion Battery Defects Classification Using Few-shot Learning Approach

Ke Wu¹, Jie Tan²*

¹ Institute of Automation, Chinese Academy of Sciences, Beijing, China
² School of artificial intelligence, University of Chinese Academy of Sciences, Beijing, China
*Corresponding author’s e-mail: jie.tan@tom.com

Abstract. In this paper, we propose a triplet siamese model for lithium-ion battery defects classification. It is a difficult task to detect the surface defects of lithium-ion batteries with stainless steel surface. The lack of three-dimensional information and the lack of marker datasets due to reflections prevent two-dimensional computer vision detection methods from meeting classification needs. In this work, the multiple exposure structured light method is utilized to obtain the three-dimensional shape of a lithium-ion battery with a stainless steel surface. The defect point cloud with three-dimensional information is obtained by this method, and then the 3D information of the defect point cloud is converted into grayscale information, and the grayscale image is used as the target domain data of the triplet siamese network. The public dataset MiniImageNet is utilized as the training data of the triplet siamese network model. The accuracies of the experimental results are 88.9%, 95.6%, and 97.8% for 1-shot, 5-shot, and 10-shot respectively. This result proves that our method can be used for lithium-ion battery defect detection.

1. Introduction

As an essential electronic energy storage device, lithium batteries are not only a powerful weapon against the problems of air pollution and rising oil prices but also a partner in our daily life[1-2]. Lithium-ion batteries have a wide range of applications, for example consumer electronics, electric transportation, and so on. Therefore, the safety of lithium-ion batteries needs extra attention. Detecting defects on the surface of lithium-ion batteries is an important way to ensure their safety and stability [3-4].

Currently, there are mainly three major difficulties in detecting surface defects of lithium-ion batteries with stainless steel surfaces, firstly, the stainless steel surface affects image acquisition, making the application of computer vision detection methods more difficult; secondly, the lack of three-dimensional information on the surface of lithium-ion batteries makes it impossible to identify the type of lithium-ion battery surface defects by two-dimensional computer vision methods; Finally, the huge number of anomalous samples is difficult to obtain, which makes it difficult to train a deep neural network model with strong robustness.

The first two problems can be solved by using the structured light method with multiple exposure sequences. Firstly, the structured light method can acquire the 3D information of the target, and then the multiple exposure sequence method can eliminate the influence of reflected light on image acquisition. While the shortage of anomalous samples is a difficult problem to solve. In many practical
scenarios, especially in the industrial field, it is often encountered that labeled samples are difficult to be acquired in large quantities. Unlike humans, deep neural networks are difficult to learn new knowledge with relatively few learning samples. Recently a field named the few-shot learning (FSL) has been demonstrated [5-7]. How to learn new classes on deep neural network models using a small number of samples has become the main goal of research in this field.

To solve the problem of training deep network models with few shots, different approaches are proposed. For example, Dandage, H. K. proposed a multi-scale image method [8]. This method is similar to the Siamese network approach, which designs a skillful structure to take a pair of image devices as input samples, and then training the network model so that the input samples of different categories are as far away as possible in the first stage, and finally separating the samples of two categories in the second stage. A similar approach to solve the few-sample problem is the Siamese network. This model has been applied to solve the few-shot problem in the field of plant disease, PCB detection, and so on [9-10]. Besides that, data augmentation and weak-supervised learning are also can be used to deal with few-shot learning problems [11-12].

In order to deal with the above problems. In this paper, a triplet Siamese network model for the lithium-ion battery defects classification approach is proposed. First, a phase shift method based on multiple exposures is used to obtain accurate three-dimensional information of the Li-ion battery. After that, the defective part of the anomalous Li-ion cells can be separated by aligning with the criteria model. Finally, defective Li-ion cells can be well classified by triple Siamese network.

2. Approach
The Triplet Siamese network for lithium-ion battery defects classification method proposed in this paper consists of two main parts, firstly, a multi-exposure structured light 3D reconstruction method, and secondly, a lithium-ion battery defect classification model based on the Siamese network. Figure 1 shows an overview of the proposed approach, described as follows.

2.1. Multiple exposure fusion

2.1.1. Digital fringe projection technique with phase shift
Generally, a classic fringe pattern of the digital fringe projection technique with phase shift method can be described as follows [13].

$$L(u,v) = I^*(u,v) + I(u,v) \cos[\varphi(x,y) + \theta_k]$$

(1)

Where $u$ and $v$ represents projector's pixel horizontal and vertical coordinates respectively. $I^*(u,v)$ is the intercept value, $I(u,v)$ is the amplitude, $\varphi$ represents the phase value that we want to acquire, and $\theta_k$ represents the phase shift. In general, we set $\theta_k = 2\pi(k-1)/3$. The phase $\varphi$ can be obtained by following equation:

$$\varphi(u,v) = \arctan\left\{\frac{\sqrt{3}[L_1(u,v) - L_3(u,v)]}{2L_2(u,v) - L_1(u,v) - L_3(u,v)}\right\}$$

(2)

By acquiring the phase unwrapping method and the wrapped phase $\varphi(x,y)$, continuous phase map can be computed from the unwrapped phase as mentioned in [14].

2.1.2. Exposure fusion
Exposure fusion is an image fusion method. Generally speaking, 3 main quality indicators need to be calculated: contrast, saturation and exposure of the image. While reconstructing the three-dimensional shape of the target, the method is used in the case where the target has a high dynamic range reflectivity surface. Then, the cameras are utilized to capture all images under different exposures.
Finally, by collapsing the input images using the weighted blending method, the fused results can be obtained. Please see [15] for detail.

2.1.3. Height-grey transformation

Assuming that $P_0$ is the prepared criteria three-dimensional cloud points. The three-dimensional reconstruction result produced by multiple exposure fusion methods is defined as a three-dimensional cloud point $P$. The disparity three-dimensional cloud point $P_d$ can be obtained after aligning $P$ to $P_0$, which a height threshold $a_0$ in the 3rd dimension. Then, the two-dimensional image $I_x$ can be acquired by transforming $P_d$ into an 8-bit grey image.

Lithium-ion batteries have three main defects: dents, bumps and scratches, which need to be detected in our tasks. $a_0$ is related to the producing process, the grayscale point cloud with no height information exceeding the threshold value indicates that the inspection object corresponding to the point cloud is a product with qualified appearance, or it indicates that the object is a defective product.

The height-grey transformer can be conducted as:

$$I_p(u,v) = 255 \frac{P(u,v) - P_0(u,v)}{H}$$  \hspace{1cm} (3)

Where $H$ is the farthest distance between a pair of points in all 3D point cloud pairs.

![Figure 1. Schematics of the proposed Triplet Siamese network model.](image)

### 2.2. Triplet Siamese network model with batch spectral regularization

#### 2.2.1. Model

As shown in Figure 1, the whole lithium-ion battery defect classification process is divided into two stages: pre-training and fine-tuning. In the pre-training stage, we mainly use the Triplet Siamese network model, which is a variation of the Siamese network model, similar to Siamese network, in
which all sub-networks share the same weights in the Triplet Siamese network model. When training the model, the triplet $X_p$, $X_a$, and $X_n$ are fed into the three sub-networks of the Triplet network model as the input images. In this method, regular CNN model $F_\varphi$ is used to obtain image features, such as Resnet, VGG, etc [16-17]. The public dataset is used as source domain data for pre-training the triplet model. In this stage, the weight parameters of the feature extractor, which is able to extract advanced features $f$, and the weight parameters of the fully connected layers $FC$ are trained and updated. Then Triplet loss $Loss$ is used to ensure that features of different categories can be separated in the embedding space, while features of similar categories can be as close as possible in the embedding space.

In the fine-tuning stage, the height-grayscale transformed result from the multi-exposure fusion phase shift method are trained as the training set for the classification model. In this stage, the feature extractors $F_\varphi$ from the training phase are directly adopted. Therefore, this method is to achieve knowledge migration from large to small datasets by fixing the weight parameters of the CNN image feature extractor $F_\varphi$ trained in the pre-training environment and adjusting the parameters of the CNN classifier $C_1$ by fine-tuning.

2.2.2. Objective Function

The objective function reflects the loss between the estimated value and the true value in a certain mapping space, and a good objective function can help Siamese networks generate more easily distinguishable representations. Generally, the objective function contains a loss term part and a regularization term part. In face recognition tasks, it is necessary to determine whether two faces are the same. The contrastive loss used in the Siamese network calculates the similarity between two samples, but only the similarity is not enough. If we want to get a better model, we not only need to distinguish the positive and negative samples but also make the distance between the intra-class samples smaller and the inter-class samples larger. This requires the usage of Triplet loss, which is trained with two positive examples and one negative example. Usually, the triplet loss can be designed as follows:

$$L_t = \max \left\{d \left( X_a, X_p \right) - d \left( X_a, X_n \right) + \text{margin,0} \right\}$$  \hspace{1cm} (4)

Where $X_a$ denotes the anchor image, $X_p$ denotes the positive image, and $X_n$ denotes the negative image. margin is a hyper-parameter that indicates how far $d \left( X_a, X_p \right)$ should be from $d \left( X_a, X_n \right)$, e.g. margin = 0.4 margin = 0.4 , $d \left( X_a, X_p \right) = 0.6$, then $d \left( X_a, X_n \right)$ should be greater than 1.0.

To improve the generalization performance of the model migration model, we add a batch spectral regularization (BSR) term to the loss function. Bing and Chen had experimentally verified that the BSR term has some performance in improving the generalization ability of the model [18-19].

Supposing that $\{X_i, X_j\}_{j=1,j=1}^{b}$ is a batch of training data, and $f_i = F_\varphi (X_i)$ is the feature vector for the i-th instance in the batch. The feature matrix can be written as $M = [f_1, f_2, \cdots, f_b]$. The BSR term can be calculated as:

$$L_{BSR} = \sum_{i=1}^{b} \sigma_i^2$$  \hspace{1cm} (5)

Where are $\sigma$ singular values of the batch feature matrix $M$. And the final loss of the Triplet Siamese network can be written as:

$$Loss = L_t + \lambda L_{BSR}$$  \hspace{1cm} (6)
Where
\[ L_{\text{BSR}} = l_{\text{BSR}1} + l_{\text{BSR}2} + l_{\text{BSR}3} \cdot l_{\text{BSR}1}, l_{\text{BSR}1}, l_{\text{BSR}1} \]
are three subnet BRS value of the triplet Siamese network.

2.2.3. CNN Classifier.
As shown in Figure. 1. We used a shallow CNN \( C_i \) of dimension 1024 to 3 as a classifier for classifying the distance between the features of different classes in the fine-tuning stage.

3. Experiments

3.1 Setup
In this experiment, the source domain samples are obtained from the MinImage public dataset, which contains a total of 60,000 samples and 6000 categories, and the dataset is divided into a training set (80%) and a test set (20%) for training the Triplet Siamese network. The target domain samples contain a total of 60 samples and 3 lithium-ion battery defect classes. For each class, 1, 5, and 10 images are randomly selected to form the target domain training set, and the remaining data set as the test set.

For the experimental parameters, when pre-training the model in the source domain, the values of training epoch, learning rate, batch size, \( \lambda \), and margin are 1000, 1e-7, 1e-4, 64, and 1.0, respectively. When fine-tuning the classifier with the target domain dataset, the gradient descent optimization algorithm is used, and the values of momentum, learning rate, and weight decay are set to 0.9, 1e-2, and 1e-3, respectively. The training epoch is set as 300.

3.2 Result
As shown in Table 1, the accuracies of the proposed lithium-ion battery defect classification model in the target domain are 0.889%, 0.956%, and 0.978% for 1-shot, 5-shot and 10-shot respectively, while the classification accuracies without the BSR term are 0.844%, 0.933%, and 0.933% for 1-shot, 5-shot and 10-shot respectively. As a comparison, the classification accuracy of the Siamese network using contrastive loss and the Siamese network with the BSR term are 0.867%, 0.933%, 0.911%, and 0.733%, 0.867%, 0.889%, respectively. The results show that the Triplet Siamese network proposed in this method has a better ability to learn new defect classes of lithium-ion batteries and can better classify different defects.

| Methods         | 3way1shot | 3way5shot | 3way10shot |
|-----------------|-----------|-----------|------------|
| Siamese         | 0.733     | 0.867     | 0.889      |
| Siamese + BSR   | 0.867     | 0.933     | 0.911      |
| Triplet Siamese | 0.844     | 0.933     | 0.933      |
| Triplet Siamese + BSR | **0.889** | **0.956** | **0.978** |

4. Conclusion
In this work, a triplet Siamese network approach for the classification of the Li-ion battery using a few-shot learning method is proposed. The experiment result indicate that our approach can overcome the effect of reflectivity, the limitation of acquiring the 3D information, and the lack of large training data. The accuracy of the target domain is 88.9%, 95.6%, and 97.8% respectively for 1-shot, 5-shot and 10-shot, which shows that the proposed method is useful for the new defects class classification of Li-ion batteries with a few new labelled data.
Acknowledgments
This work was supported in part by the National Natural Science Foundation of China under Grants U1801263 and U1701262.

References
[1] Ali, M. U., Zafar, A., Nengroo, S. H., Hussain, S., Park, G. S., & Kim, H. J. (2019). Online remaining useful life prediction for lithium-ion batteries using partial discharge data features. Energies, 12(22), 4366.
[2] Lukatskaya, M. R., Dunn, B., & Gogotsi, Y. (2016). Multidimensional materials and device architectures for future hybrid energy storage. Nature communications, 7(1), 1-13.
[3] Bruce, P. G., Scrosati, B., & Tarascon, J. M. (2008). Nanomaterials for rechargeable lithium batteries. Angewandte Chemie International Edition, 47(16), 2930-2946.
[4] Nitta, N., Wu, F., Lee, J. T., & Yushin, G. (2015). Li-ion battery materials: present and future. Materials today, 18(5), 252-264.
[5] Ji, Z., Zou, X., Huang, T., & Wu, S. (2020). Unsupervised few-shot feature learning via self-supervised training. Frontiers in computational neuroscience, 14.
[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, 30, 1318-1331.
[8] Dandage, H. K., Lin, K. M., Lin, H. H., Chen, Y. J., & Tseng, K. S. (2021). Surface defect detection of cylindrical lithium-ion battery by multiscale image augmentation and classification. International Journal of Modern Physics B, 2140011.
[9] Argüeso, D., Picon, A., Irusta, U., Medela, A., San-Emerio, 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.
[10] Tang, S., He, F., Huang, X., & Yang, J. (2019). Online PCB defect detector on a new PCB defect dataset. arXiv preprint arXiv:1902.06197.
[11] Tao, X., Wang, Z., Zhang, Z., Zhang, D., Xu, D., Gong, X., & Zhang, L. (2018). Wire defect recognition of spring-wire socket using multitask convolutional neural networks. IEEE Transactions on Components, Packaging and Manufacturing Technology, 8(4), 689-698.
[12] Marino, S., Beauséroy, P., & Smolarz, A. (2019). Weakly-supervised learning approach for potato defects segmentation. Engineering Applications of Artificial Intelligence, 85, 337-346.
[13] Ekstrand, L., & Zhang, S. (2011). Autoexposure for three-dimensional shape measurement using a digital-light-processing projector. Optical Engineering, 50(12), 123603.
[14] Zhang, S. (2018). Absolute phase retrieval methods for digital fringe projection profilometry: A review. Optics and Lasers in Engineering, 107, 28-37.
[15] Wu, K., Tan, J., Xia, H. L., & Liu, C. B. (2020). An exposure fusion-based structured light approach for the 3D measurement of a specular surface. IEEE Sensors Journal, 21(5), 6314-6324.
[16] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
[17] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
[18] Chen, X., Wang, S., Long, M., & Wang, J. (2019, May). Transferability vs. discriminability: Batch spectral penalization for adversarial domain adaptation. In International conference on machine learning (pp. 1081-1090). PMLR.
[19] Liu, B., Zhao, Z., Li, Z., Jiang, J., Guo, Y., & Ye, J. (2020). Feature transformation ensemble model with batch spectral regularization for cross-domain few-shot classification. arXiv preprint arXiv:2005.08463.