Rotor Winding Image Detection Method Based on Model-based Transfer Learning

Jia Youbin1, Chen Gang1, Zhang Xiaoguo2\*

1. *Department of Mechanical Engineering*
2. *Department of Instrument Science and Engineering*

*University of Southeast*

*Nanjing, jiangsu Province, China*

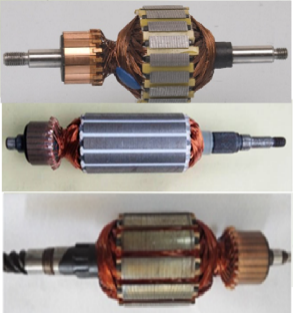
*xgzhang@seu.edu.cn*

Abstract—Rotor is a core component of the motor. The winding qualification of rotor is one of the core factors for the proper functioning of the rotor. However, the detection of the qualification of the winding is still carried out by manual operation, which is prone to cause false detection and low efficiency. Hence, it is important to achieve automatic detection of the rotor windings and enhance the detection accuracy. But the challenge is that winding image dataset of different types of rotor exist dataset bias and there are limited labeled examples of rotor winding image. To resolve the dataset bias problem, we proposed a new image binarization method to binarize winding image. Using the binarized image to train and test model can significantly reduce the interference of dataset bias. Meanwhile, we proposed a method to build model-based transfer learning model which is based on the pre-trained Inception-V3 model trained with the ImageNet dataset. The model is used to resolve the problem of labeled examples are limited. The comparing experiments show that, the model-based transfer learning model training and testing with binarized image significantly outperform all other models, and can achieve stable and accurate detection of the rotor image.

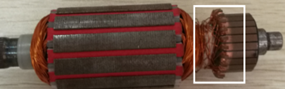
Keywords—qualification detection; rotor winding; binarized image; Inception-V3; model-based transfer learning

# Introduction

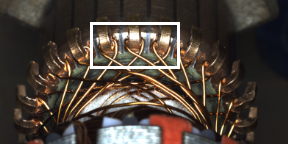
As a core product of industrial automation, motor is widely used in Large-scale Electromechanical equipment, elevators, refrigerators, air conditioning and other places in the fields of industry and agriculture. Therefore, the demand for motors increases daily. As the core component of the motor, the quality of the rotor winding (see Fig. 1) is critical to the service time of the motor. Based on current rotor production technology, rotor windings are prone to occur missing winding, broken winding and so on (see Fig. 2). However, the current method for detected the qualification of rotor windings on the assembly line by the manual work. Considering of the complex production environment, this method has low efficiency and with the high cost of the working time, the manual work easily occurs misjudgement, and is not suitable for mass production. Therefore, it is very urgent to find a new alternative detection method that can quickly detect the qualification of the rotor windings.

(a) Different types of rotor



(b) Winding area

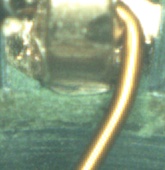


(c) Camera field of view

Figure 1. Different types of rotor (a); The position of the winding: Winding area to be detected on the rotor (b); The position of the winding to be detected in the camera field of view (c).

The concept of defect detection based on machine vision and machine learning was popularized in recent years [1]. We recognize that there is dataset bias in different types object or under different lighting conditions. The impact of the dataset bias on detection results is enormous.

Recently, convolutional neural network (CNN) based feature representations have been proved to be extremely effective for a

(a) Missing winding (b) Broken winding (c) Qualified winding

Figure 2. Different forms of winding: Rotor winding is missed in this position (a); Rotor winding is broken in this position (b); Rotor winding is qualified in this position (c).

variety of visual recognition tasks [2,3, 4,5]. In particular, using deep representations dramatically reduce the effect of resolution and lighting on domain shifts [2,6]. However, training an effective CNN model to recognize rotor winding image requires a large number of training samples, but there are limited labeled examples of rotor winding.

As the improvement method, training a joint CNN architecture was proposed [7], but has only two layers. The deeper architecture [4] outperform the architecture with only two layers. A new CNN architecture [8] which introduces an adaptation layer and an additional domain confusion loss, but was limited to only one adaptation layer, and was outperformed by the architecture which has three adaptation layers [9]. However, these methods require similar dataset between source and target domain and therefore it is difficult to extend these methods to the domains which belong to unrelated domains.

As for as the rotor winding image detection concerned, the labeled training samples are limited, directly train convolutional neural network on a small amount of labeled examples turns out to be problematic. Fortunately, pre-trained CNN do perform well in new domains. And using the deep mid-level features learned on ImageNet, instead of the more conventional bag-of-words features, effectively removed the bias in some of the domain adaptation settings [2,10]. In fact, the dataset bias is often eliminated by learning a feature space transformation to align the source and target representations [11,12,13,14].

We propose a transfer learning method to detect the rotor winding, which uses a pre-trained Inception-V3 model [15] trained with ImageNet to build a new model. The new model is trained and tested with the binarized images which is binarized directly from color image by the features of RGB value. Finally, an accurate detection classifier is obtained. The experiment analysis demonstrating that the method is significantly outperform all other methods.

# Model-based Transfer learning

We introduce a transfer learning method which we use to learn an image classifier that is both stabilized and which offers strong adaptability. As shown in the Fig. 3, on the one hand, we begin with an Inception-V3 model, and pre training the Inception-V3 model with ImageNet dataset, then save the Inception-V3 model. On the other hand, we use automatic equipment to collect the rotor winding image, then we binarize



Figure 3. Technological process of model-based transfer learning

the rotor winding image and constructing training dataset and different testing dataset using the banirized image. In order to transfer the Inception-V3 model from ImageNet dataset to rotor winding dataset, we use training dataset constructed by binarized image to fine-tuning the fully connected layer of pre-trained Inception-V3, and then save the Inception-V3 model. Finally, we use different testing dataset constructed by binarized image to test the final model. It has been shown that a pre-trained Inception-V3 model can be adapted to different domains through model-based transfer learning.

## Image banarization based on RGB features

Model-based transfer learning is challenging in that there is dataset bias of winding image dataset between different rotor. To resolve this problem, our intuition is that if we can binarize the winding image dataset, therefore, we binarized the color image by the features of RGB value in order to simplify the processing of neural network and highlight the features of winding. The bianrized image extract the winding from the background, eliminating the effects of background and lighting conditions and reducing the dataset bias.

This paper mainly focuses on the qualification of the winding. The qualification of winding is detected by the detection algorithm. For the current production conditions, there are some non-conforming winding between the rotor commutator and the copper wire. such as missing winding, broken winding, which are shown in Fig. 2.

In order to detect the non-conforming winding, we need to highlight the features of the winding. Our intuition is that the winding image (see Fig.4 (a)) needs to be binarized to reduce the influence the dataset bias. Generally, the image binarization method needs to grayscale, filter and thresholds the image. The winding in the image is target, while the other unrelated areas are the background, and an optimal threshold needs to be determined to distinguish the two parts. Considering that the gray scale distribution between the winding and the background is similar in some parts, the brightness of the winding and the hook background change with the light irradiation angle, so it is difficult to use a fixed gray threshold to extract winding from the image.

Normally, the adaptive algorithm [16] is used to adaptively calculate the threshold. and the binarized image obtained after the grayscale, filtering and threshold. The following figure shows a binarized winding image was bianarized by the adaptive method (see Fig.4 (b)). It could be found that there is some interference in the background area above the adaptive threshold image, and the obtained winding area is incomplete. Making further efforts, we use OTSU algorithm [17] to binarize the image (see Fig.4 (c)). It could be found that the obtained winding area is more complete, but there is still some interference in the background area.

To resolve this problem, we propose a new binarization method. After analyzing the RGB values of the color image, we found that the color of the winding is yellow, and the RGB value features of yellow winding is different from the background, which is characterized by , where represents the value of the R channel, represents the value of the G channel, and represents the value of the B channel. Therefore, the new image binarization algorithm is that for each pixel in color winding image,

(1)

Then we obtain a new binarized image (see Fig. 4(d)) which has three channel and the winding part is complete and the features are significant.

    
 (a) Color image (b) Adaptive

(c) OTSU (d) RGB

Figure 4. Image binaryzation：(a) color image of rotor winding; (b) using adaptive threshold to binarize image; (c) using OTSU method to binarize image; (d) the new image binaryzation method which is based on the features of RGB value.

## Training model

We introduce a new method to train convolutional neural network (CNN) which we use to learn a visual representation that is both domain invariant. It has been shown that a pre-trained CNN can be adapted for a new task through fine-tuning. We begin with the Inception-V3, which has five convolutional layers, one convolutional padded layers, two pool layers, three Inception layers, one linear layers and softmax classifier. The outline of the proposed network architecture of Inception-V3 is shown in the following table 1.

The reason we choose Inception-V3 model is that the Inception-V3 have provided several design principles to scale up convolutional networks. This guidance can lead to high performance vision networks that have a relatively modest computation cost compared to simpler, more monolithic architectures. This might prove to be helpful in systems for detecting relatively small objects. [15].

Our intuition is that a lower dimensional layer can be used to regularize the training of the source classifier and prevent overfitting to the particular nuances of the source distribution. We froze lower layers before linear layer, and we fine-tuning the fully connected layers after the frozen layers.

In this paper, we use binarized image to train and test the model. But the main challenge is that the target domain has just limited labeled samples. If we use these samples directly to train the model which is prone to problematic.

In order to resolve the problem, the ImageNet is used for training the parameters of the frozen layers to learn a representation that is domain invariant, the source labeled dataset of the rotor winding is used for training the parameters of the fully layer to learn a classifier that is able to recognize rotor winding image. Then we save the model to test different dataset of rotor winding. We expect that such a processing of model-based transfer learning will thus enable increased adaption performance.

Table 1. The outline of the proposed network architecture of Inception-V3

|  |  |  |  |
| --- | --- | --- | --- |
| type | patch size/stride  or remarks | input size | Train/ frozen |
| conv | 3×3/2 | 299×299×3 | frozen |
| conv | 3×3/1 | 149×149×32 | frozen |
| conv padded | 3×3/1 | 147×147×32 | frozen |
| pool | 3×3/2 | 147×147×64 | frozen |
| conv | 3×3/1 | 73×73×64 | frozen |
| conv | 3×3/2 | 71×71×80 | frozen |
| conv | 3×3/1 | 35×35×192 | frozen |
| 3×Inception | as in paper[15] | 35×35×288 | frozen |
| 5×Inception | as in paper[15] | 17×17×768 | frozen |
| 2×Inception | as in paper[15] | 8×8×1280 | frozen |
| pool | 8×8 | 8×8×2048 | frozen |
| linear | logits | 1×1×2048 | train |
| softmax | classifier | 1×1×1000 | train |

# Experiments

We compare the state-of-the-art transfer learning based on Inception-V3 to deep learning methods based on Inception-V3 trained and tested on both binarized image and color image, focusing on the efficacy of model-based transfer learning model training and testing with binarized image.

## Setup

In this paper, the detection method need a pre-trained model. In order to satisfy the required transferability, we use ImageNet to pre train the Ineption-V3 model.

**ImageNet** ImageNet is an image dataset organized according to the WordNet hierarchy. Each meaningful concept in WordNet, possibly described by multiple words or word phrases, is called a "synonym set" or "synset". There are more than 100,000 synsets in WordNet, majority of them are nouns (80,000+). ImageNet aim to provide on average 1000 images to illustrate each synset. Images of each concept are quality-controlled and human-annotated. In its completion, ImageNet will offer tens of millions of cleanly sorted images for most of the concepts in the WordNet hierarchy.

We begin with a pre-trained Inception-V3 model which is used for model-based transfer learning, the ImageNet is used for training the parameters of the frozen layers, the source labeled dataset of the rotor winding is used for training the parameters of the fully connected layer to learn a classifier. Then we froze the parameters and save the model. We evaluate our model across 4 domains, same type of rotor under same lighting condition, same type of rotor under different lighting conditions, different types of rotor under same lighting condition and different types of rotor under different lighting conditions (see Fig. 5).

We compare to a variety of methods: Inception-V3 training and testing with color image, Inception-V3 training and testing with binarized image, transfer Inception-V3 training and testing with color image. The number of labeled examples of training dataset is 2400. Testing dataset is divided into 4 categories: same type of rotor under same lighting condition, same type of rotor under different lighting conditions, different types of rotor under same lighting condition and different types of rotor under different lighting conditions. The number of testing examples of each category is 10000.

We compare the classification accuracy for each domain. For deep learning methods, we follow the standard procedures for model training as explained in their respective papers. For transfer learning method, we use a pre-trained model to training a new model. We use the pre-trained architecture, however, due to limited training examples in our dataset, we fix all layers before the Linear layer that were copied from pre-trained model, and fine-tuning the Linear layer, both via back propagation. We use stochastic gradient descent (SGD) with 0.01 learning rate.

## Result and Discussion

The test result on the different domains are shown in Fig. 5. We can observe that model-based transfer learning training and testing with binarized image significantly outperforms the comparison methods on most domains, and achieves comparable performance on the domain of different types of rotor under different lighting conditions. This is reasonable as the adaptability of our method across different types of rotor under different lighting conditions. The performance boost demonstrates that our method of model-based transfer learning training and testing with binarized image is able to transfer pre-trained deep models across different tasks.

Feature 5. Accuracy of different models on different testing dataset

From the experimental results, we can draw the following conclusion. (1) Model-based transfer learning methods significantly outperform conventional deep learning based methods by a larger margin. (2) Among the model-based transfer learning methods, the model training and testing with binarized image provides great improvement, suggesting that the model training and testing with binarized image can adapted to the change both rotor type and lighting condition very well. (3) When it comes to the situation that we have limited labeled examples and there is dataset bias between source dataset and target dataset. Model-based transfer learning training and testing with binarized image significantly outperform all other models.

# Conclusion

In this paper, we presented a model-based transfer learning method for classification. This method makes use of pre-trained   
Inception-V3 trained with ImageNet and then fine-tuning on binarized image to ensure that domains are indistinguishable in the learned representation.

And we presented a new effective method how to convert color image into binarized image, which converted training dataset or testing dataset into binarized image during training or testing our model, achieve state-of-the-art performance on testing dataset which belong to different rotor types under different lighting conditions, beating other methods by a considerable margin.

Experiment results show that using pre-trained Inception-V3 trained with ImageNet to conduct transfer learning is an effective way to resolve the problem that the target domain has just limited labeled samples, and converting training and testing dataset into binarized image will significantly enhance the adaptability of the model.

REFERENCES

1. Huang Kaiqi, Ren Weiqiang, and Tan Tieiu. “Summary of Image Object Classification and Detection Algorithms,” Chinese Journal of Computers ,37(06):pp. 1225-1240, 2014.
2. J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. “DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition,” In Proc.ICML, 2014.
3. R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich. “feature hierarchies for accurate object detection and semantic segmentation,” arXiv e-prints, 2013.
4. A. Krizhevsky, I. Sutskever, and G. E. Hinton. “ImageNet classification with deep convolutional neural networks,” In Proc. NIPS, 2012.
5. P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun. “Overfeat: Integrated recognition, localization and detection using convolutional networks,” CoRR, abs/1312.6229, 2013.
6. J. Hoffman, E. Tzeng, J. Donahue, , Y. Jia, K. Saenko, and T. Darrell. “One-shot learning of supervised deep convolutional models,” In arXiv 1312.6204; presented at ICLR Workshop, 2014.
7. S. Chopra, S. Balakrishnan, and R. Gopalan. “DLID: Deep learning for domain adaptation by interpolating between domains,” In ICML Workshop on Challenges in Representation Learning, 2013.
8. Tzeng, E., Hoffman, J., Zhang, N., Saenko, K., and Darrell, T. “Deep domain confusion: Maximizing for domain invariance,” Technical report, arXiv:1412.3474, 2014.
9. M. Long and J. Wang. “Learning transferable features with deep adaptation networks,” CoRR, abs/1502.02791, 2015.
10. Y. Ganin and V. Lempitsky. Unsupervised Domain Adaptation by Backpropagation. ArXiv e-prints, Sept. 2014.
11. K. Saenko, B. Kulis, M. Fritz, and T. Darrell. Adapting visual category models to new domains. In Proc. ECCV, 2010.
12. B. Kulis, K. Saenko, and T. Darrell. What you saw is not what you get: Domain adaptation using asymmetric kernel transforms. In Proc. CVPR, 2011.
13. B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars. Unsupervised visual domain adaptation using subspace alignment. In Proc. ICCV, 2013.
14. B. Gong, Y. Shi, F. Sha, and K. Grauman. Geodesic flow kernel for unsupervised domain adaptation. In Proc. CVPR, 2012.
15. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. “Rethinking the inception architecture for computer vision,” arXiv preprint arXiv:1512.00567, 2015.

[16] Mao Xingyun，Long Xuefei, Opencv3 Programming Introduction，Electronics Industry, 2015.

[17] Otsu N. A threshold selection method from gray-level histogram. IEEE Trans,1979;SMC-9;62-66.