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

Youbin Jia and Gang Chen\*  Xiaoguo Zhang

*Department of Mechanical Engineering Department of Instrument Science and Engineering*

*University of Southeast University of Southeast*

*Nanjing, jiangsu Province, China Nanjing, jiangsu Province, China*

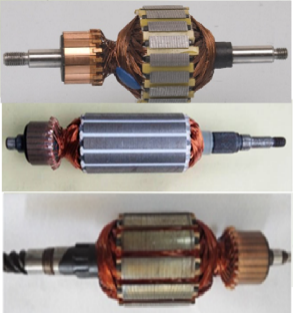
*ldd1485@163.com xgzhang@seu.edu.cn*

Abstract—Rotor is a core component of the electric motor. The qualification of rotor winding is one of the core factors for the proper functioning of the rotor. However, the qualification detection of the winding is still carried out by manual operation, which is prone to cause miscalculation and low efficiency. Hence, it is important to achieve automatic detection of the rotor windings and enhance the detection accuracy. Recently, convolutional neural network (CNN) has been successfully applied to image recognition, but to achieve high accuracy requires a large number of labeled samples and there is almost no dataset bias between the target dataset and the source dataset. But the challenges of using CNN to recognize rotor winding are that the winding image dataset of different types of rotor exist large dataset bias and the labeled examples are limited. Tosolve the dataset bias problem, we proposed a new image binarization method to get binary rotor winding images. Using the binary images 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 solve the problem that labeled samples are limited. The comparing experiments show that the model-based transfer learning model trained and tested with binary images significantly outperform all other models, and can achieve stable and accurate detection of the rotor images.

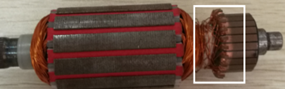
Keywords—qualification detection; rotor winding; binary image; Inception-V3; model-based transfer learning

# Introduction

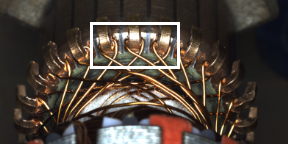
As an essential product of industrial automation, electric motor is widely used in large-scale electromechanical equipment, elevators, refrigerators, air conditioning and other places in the fields of industry and agriculture, and the demand for electric motors increases daily. As the core component of the electric motor, the qualification of the rotor winding (see Fig. 1) is critical to the service life of the electric motor. According to current rotor production technology, rotor windings are prone to loss，fracture and other situations (see Fig. 2). However, the qualification of winding is still detected by the manual work. Considering of the complex production environment, this method not only has low efficiency and high time cost, but also easily occurs misjudgements 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 rotor winding.

(a) Different types of rotor



(b) Winding area

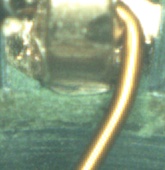


(c) Camera field of view

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 has become more and more popular recently [1]. We recognize that there is dataset bias in different types of objects or objects under different lighting conditions. The impact of the dataset bias on detection results is enormous.

Recently, the convolutional neural network (CNN) based feature representation has 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 can dramatically reduce the effect of resolution and lighting on domain shifts [2,6]. However, training an effective CNN model to recognize rotor winding images requires a large number of training samples, but there are limited labeled samples of rotor winding.

As an improvement method, training a joint CNN architecture with two layers was proposed [7]. The deeper architecture [4] outperforms the architecture with only two layers. A new CNN architecture [8] introduces an adaptation layer and an additional domain confusion loss but is limited to only one adaptation layer and 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 far as the rotor winding image detection concerned, the labeled samples are limited, and directly training convolutional neural network on a small amount of labeled samples 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 removes 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].

Dataset bias appears on many tasks, and according to the characteristics of the rotor winding images, we can eliminate the dataset bias of the rotor winding images by extracting the winding features from the image. Normally, the binary image processing is an effective method to reduce the influence of background and lighting conditions. In addition, using the deep mid-level features learned on ImageNet, effectively removes the bias in some of the domain adaptation settings and solves the problem that labeled samples are limited [2,6]. These principles form the essence of our proposed approach.

We propose a new 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 binary images which is transformed directly from color images by the features of RGB values. Finally, an accurate detection classifier is obtained. According to the experiment, the method is significantly outperform all other methods.

# pre-trained inception-v3 model trained and tested with binarized image

We introduce a transfer learning method to build an image classifier that has both stability and adaptability. As shown in the Fig. 3, on the one hand, we begin with an Inception-V3 model, and pre-train 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 images, then we binarize the rotor winding images and construct training dataset and different testing dataset using the binary images. In order to transfer the Inception-V3 model from ImageNet dataset to rotor winding dataset, we use training dataset constructed by binary images to fine-tune the fully connected layer of pre-trained Inception-V3, and then save the Inception-V3 model. Finally, we use different testing dataset constructed by binary images to test the final model. It shows that a pre-trained Inception-V3 model can be adapted to different domains through model-based transfer learning.

## Image banarization based on RGB features

One of the challenges of using CNN to recognize rotor winding is that the winding image dataset exist bias. To solve this problem, we binarize the color images by the features of RGB values to simplify the processing of neural network and highlight the features of winding. The   
  


Figure 3. Technological process of model-based transfer learning  
  
binary images extract the winding from the background, eliminate the effects of background and lighting conditions and reduce 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 and broken winding, which are shown in Fig. 2.

In order to detect the non-conforming winding, the winding images (see Fig.4 (a)) need to be binarized to reduce the influence of background and lighting conditions. The winding in the image is the 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 areas, 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.

To solve 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.

For each pixel in the color winding image, the RGB values are binarized by the assignment function

(1)

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

In order to confirm the superiority of our method, the adaptive algorithm [16] and the OTSU algorithm [17] are used for comparison. The Fig.4 (b) shows a binary winding image which was binarized by the adaptive algorithm. 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. The Fig.4 (c) shows a binary winding image which was bianarized by the OTSU algorithm. It could be found that the obtained winding area is more complete, but there is still some interference in the background area. Apparently, the new binarization algorithm outperform the adaptive algorithm and the OTSU algorithm.

## Training model

We introduce a new method to train convolutional neural network (CNN) to obtain a visual representation that is domain invariant and we choose the Inception-V3 architecture as the training model. We first give an overview of the Inception-V3 architecture, the Inception-V3 has five convolutional layers, one convolutional padded layers, two pool layers, three Inception layers, one linear layers and a softmax classifier. The outline of the proposed network architecture of Inception-V3 is shown in the following table 1.

    
 (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.

The reason why we choose the Inception-V3 architecture is that the Inception-V3 has 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 and more monolithic architectures. This might prove to be helpful in systems for detecting relatively small objects [15].

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

In this paper, the binary images are used to train and test the model to solve the dataset bias problem. But another challenge is that the target domain has limited labeled samples. If we use these samples directly to train the model, it is prone to problematic.

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 |

In order to solve the problem, the ImageNet is used for training the parameters of the frozen layers to obtain a representation that is domain invariant, and the source labeled dataset of the rotor winding is used for training the parameters of the fully connected layer to build a classifier that is able to recognize rotor winding images. Then 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.

# Experiments

In order to confirm the superiority of our method, we compare the state-of-the-art transfer learning method with deep learning method by training and testing them on both binary images and color images, focus on the efficacy of model-based transfer learning model trained and tested with binary images.

## Setup

In this paper, the transfer learning 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, and 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.

The ImageNet is used for training the parameters of the frozen layers, and the source labeled dataset of the rotor winding is used for training the parameters of the fully connected layer to build 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).

In order to highlight the effectiveness of our proposed method, it compares to a variety of methods: Inception-V3 trained and tested with color images, Inception-V3 trained and tested with binary images, transfer Inception-V3 trained and tested with color images. The number of labeled samples 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 samples of each category is 10000.

The classification accuracy for each domain is compared. For deep learning method, 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 train a new model. We use the pre-trained architecture, however, due to limited training samples in our dataset, we fix all layers before the liner layer that were copied from pre-trained model, and fine-tune the liner layer, both via back propagation. We use stochastic gradient descent (SGD) with 0.01 learning rate.

## Result and Discussion

The test results on the different domains are shown in Fig. 5. It could be found that model-based transfer learning method trained and tested with binary images significantly outperforms the comparison methods on most domains, and achieves comparable performance on the domain of different types of rotor under different lighting conditions. It 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 trained and tested with binary images 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 method significantly outperforms conventional deep learning based method by a larger margin. (2) Among the model-based transfer learning methods, the model trained and tested with binary images provides great improvement, suggest that the model trained and tested with binary images can adapt to the changes of both rotor type and lighting condition very well. (3) When it comes to the situation that we have limited labeled samples and there is dataset bias between source dataset and target dataset. Model-based transfer learning model trained and tested with binary images significantly outperforms all other models.

# Conclusion

In this paper, we present 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 binary images to ensure that domains are indistinguishable in the obtained representations.

And we present a new effective method to convert color images into binary images. The new binarization method is used to convert training dataset and testing dataset into binarized images during training and testing our model. Using binary dataset, our model achieves state-of-the-art performance on testing dataset which belong to different types of rotor under different lighting conditions, and beats other models 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 solve the problem that the target domain just has limited labeled samples, and converting training and testing dataset into binary images 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.
18. Sinno Jialin Pan, and Qiang Yang. "A Survey on Transfer Learning." IEEE Transactions on Knowledge and Data Engineering 22.10(2010):1345-1359.