

RESEARCH

Open Access



# Automatic classification of refrigerator using doubly convolutional neural network with jointly optimized classification loss and similarity loss

Yongchang Gao<sup>1</sup>, Jian Lian<sup>2\*</sup> and Bin Gong<sup>1\*</sup>

## Abstract

Modern production lines for refrigerator take advantage of automated inspection equipment that relies on cameras. As an emerging problem, refrigerator classification based on images from its front view is potentially invaluable for industrial automation of refrigerator. However, it remains an incredibly challenging task because refrigerator is commonly viewed against dense clutter in a background. In this paper, we propose an automatic refrigerator image classification method which is based on a new architecture of convolutional neural network (CNN). It resolves the hardships in refrigerator image classification by leveraging a data-driven mechanism and jointly optimizing both classification and similarity constraints. To our best knowledge, this is probably the first time that the deep-learning architecture is applied to the field of household appliance of the refrigerator. Due to the experiments carried out using 31,247 images of 30 categories of refrigerators, our CNN architecture produces an extremely impressive accuracy of 99.96%.

**Keywords:** Image classification, Machine vision, Industry automation

## 1 Introduction

Being one of the most common household appliances in industrial production, the refrigerator plays a vital role in humans' daily lives, globally. In recent decades, China ranks the first in refrigerator production; the refrigerator industry has considerably contributed to China's economy. It has been reported by the National Bureau of Statistics of China [1] that China produced 79.92 million refrigerators in 2015.

Various vision-based applications in refrigerator manufacturing process have been brought forward along with the development of industrial automation especially in the emergence of industry 4.0 standardization [2–4]. As an example, automatic classification is potentially invaluable

for industrial automation. It helps to resolve several disadvantages of manual inspection which is being widely adopted on the production line, e.g., high labor cost and inter/intra-observer variations.

As an emerging problem, image-based automatic classification of the refrigerator is of great importance because it can potentially provide a valuable tool for industrial automation of refrigerator. There have been various machine-learning-based image classification methods proposed in industrial automation. However, none of them was designed for the automatic classification of the refrigerator, which mainly relies on human's visual system at present. The machine-learning-based classification methods can be roughly divided into two categories, including classifier combined with handcrafted features and automatically extracted features. With experiments, we respectively testify the support vector machine (SVM)-based methods [5–8] and CNN-based methods, in which the front-view image of the refrigerator is exploited due to the rich information captured by this image in recognizing and differentiating refrigerators.

\*Correspondence: [lianjianlianjian@163.com](mailto:lianjianlianjian@163.com); [gb@sdu.edu.cn](mailto:gb@sdu.edu.cn)

<sup>2</sup>Department Electrical Engineering Information Technology, Shandong University of Science and Technology, Jinan 250031, China

<sup>1</sup>Department of Computer Science and Technology, Shandong University, Jinan 250010, China

Experimental results show that the CNN-based methods outperform the state-of-the-art SVM-based methods with handcrafted features. Both SVM-based methods and CNN-based methods are effective in the classification of refrigerator front-view images. However, we observe fundamental challenges in the front-view image of the refrigerator, including the dense clutter, varying backgrounds, large homogeneous areas, sparse salient regions, and the specular reflection.

In the field of machine learning, recent advances [9–11] in deep learning have led to breakthroughs in long-standing tasks such as language translation, text recognition and image-related problems of feature extraction [12, 13], image segmentation [14–16], and image classification [17–22]. Among all these techniques, CNN is one of the most successful methods and has acquired wide applications in image classification [9]. Recently, there is an increasing trend to develop variants of deep neural networks for the tasks that were difficult to implement previously; for instance, in [23], a deep polynomial network was presented to implement the tumor classification with small ultrasound images, and the classification accuracy for breast ultrasound image is  $92.40 + 1.1\%$ . A deep-learning framework for the face attribute prediction in the wild was proposed in [24] while prediction of the face attributes in the wild is challenging due to the complex face variations. It cascades two types of CNNs, which are fine-tuned jointly with attribute tags and pre-trained differently.

In this paper, we present a useful tool for classifying refrigerator based on images taken from its front view, which is accomplished by using a novel CNN architecture adapted to the specialty of refrigerator images. Our approach is data driven and free from handcrafted image features. It leverages a training process to automatically extract multi-scale image features which combine both the local and global characteristics of the refrigerator front-view image. The mapping between the refrigerator and its corresponding category, which is learned from these features, offers a function to automatically specify the classification of a refrigerator given a new input image. The proposed CNN architecture takes triple images as input, from which a triplet loss (similarity loss) is produced. Meanwhile, the traditional convolution layer is modified into a doubly convolutional layer in our proposed CNN architecture. The details of triplet loss and the doubly convolutional layer will be fully discussed in Section 2. Our experimental results from 31,247 refrigerator front-view images of 30 categories show that the proposed CNN architecture produces an impressive classification accuracy of 99.96%, which is considerably better than conventional classification techniques.

Our work offers at least five significant contributions as listed below:

- (I) To our best knowledge, this is the first attempt to introduce deep-learning methods into the image classification of the refrigerators.
- (II) We propose a novel CNN architecture with triple input and the doubly convolutional layer which both are adapted to the characteristics of refrigerator front-view images.
- (III) Besides the softmax loss (classification loss) that was commonly used in previous CNN frameworks, we introduce a triplet loss (similarity loss) in our proposed architecture. Meanwhile, different from the original convolution layer in conventional CNN architectures, a doubly convolutional layer is introduced in the proposed CNN architecture.
- (IV) Our approach performs with an impressive superiority to both the state-of-the-art image classification techniques and human's visual inspection.

The rest of this paper is organized as follows. In Section 2, we present our proposed refrigerator image classification approach. Section 3 contains experimental results and our discussions. In Section 4, we present the conclusion.

## 2 Our method

Our method is employed to classify refrigerator according to the front-view images of the refrigerator. From the observation, we have found various challenges in the images. Aiming at solving these difficulties, we propose a novel CNN architecture as described in Section 2.2.

### 2.1 The challenges in the front-view image of refrigerator

We observed there are various challenges in the front-view images of the refrigerator including dense clutter, varying backgrounds, large homogeneous areas, subtle salient regions, and specular reflection, which are described in details as follows:

- Dense clutter. With the observation, we notice there is a dense clutter in the front-view images of the refrigerator. The paper and sellotape attached on the refrigerator at the left of Fig. 1 are typical clutter in the images of the refrigerator. Compared to the image at the right of Fig. 1, the image at the left of Fig. 1 contains redundant information from the paper and sellotape. Without appropriate processing, the clutter would cause the severe distraction to the result of classification.
- Varying backgrounds. Due to the changeable position and the surroundings of the refrigerator on the assembling line, the backgrounds in the images are different from each other. Figure 2 shows two images of refrigerators belonging to the same classification



**Fig. 1** The dense clutter in the front-view images of refrigerator

but with diverse backgrounds. As the varying backgrounds randomly append unstable factors to the images of the refrigerator, it would unavoidably increase the complexity of the image classification.

- Large homogeneous areas and subtle salient regions. There are large homogeneous areas existing in the images of refrigerator, as shown in Figs. 1, 2, 3, and 4. Meanwhile, the salient regions including the display, crack, ice maker, and handle of the refrigerator are subtle compared to the overall image, e.g., in Fig. 3

(the areas confined with the red and orange squares). Large homogeneous areas and small salient regions both would cause misclassification and greatly affect the classification accuracy.

- Specular reflection. Due to the mirror surfaces of the refrigerators, noticeable effects of specular reflection are observed in the captured images. Two refrigerators of the same classification are shown in Fig. 4; there is more apparent specular reflection shown on the surface of the left refrigerator than the



**Fig. 2** The varying backgrounds in the front-view images of refrigerator



**Fig. 3** The large homogenous areas and subtle salient regions in the front-view images of refrigerator

right refrigerator. It would produce redundant information of the surroundings in the front-view images of the refrigerator.

**2.2 The architecture of our proposed CNN**

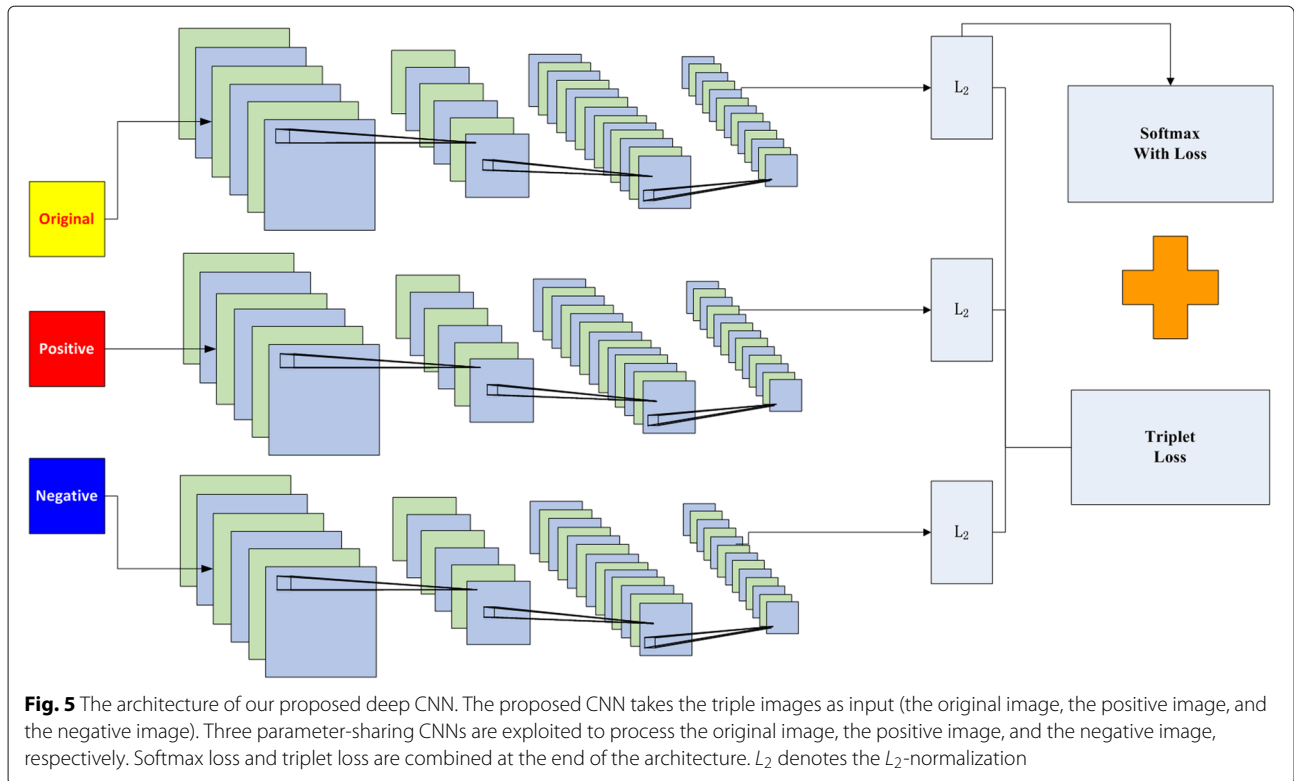
Based on the analysis of the challenges in the front-view images of refrigerators, we propose a novel CNN to perform the image classification of refrigerators. Notably, this CNN can also be employed to classify images of other household appliances, including the air conditioner,

television, and washing machine. Three convolutional layers and one fully connected layer are included in each channel of the parameter-sharing CNN. Then one softmax loss and a triplet loss are combined at the end of the architecture, which is shown in Fig. 5.

Each captured image is divided into image patches with same sizes. For each image, a positive image (same category) and a negative image (different category) are chosen from the captured images. Then, the three image patches are jointly inputted into our proposed CNN, in which



**Fig. 4** The specular reflection in the front-view images of refrigerator



three parameter-sharing CNNs are presented to handle the original, the positive, and the negative images, respectively.

Furthermore, to extend the parameter-sharing efficiency of CNN, we replace the independent convolutional filters in the convolution layers by a set of filters that are translated versions of each other, which is implemented by a two-step convolution operation or so-called doubly convolutional layer as shown in Fig. 6.

- Convolutional layer. Forty-eight kernels of  $3 \times 11 \times 11$  and  $3 \times 7 \times 7$  are applied to the input refrigerator image patch in the first layers combined with the rectified linear unit (ReLU) layer. To handle the characteristics of sparseness in refrigerator image, the operation of ReLU is performed after the convolution operation.
- Convolutional layer. One hundred twenty-eight kernels of  $3 \times 9 \times 9$  and  $3 \times 5 \times 5$  combined with the ReLU layer, the features in the input image are extracted furthermore. Same as the first layer in our CNN architecture.
- Convolutional layer. One hundred twenty-eight kernels of  $3 \times 7 \times 7$  and  $3 \times 3 \times 3$  combined with the ReLU layer, the features in the input image are extracted furthermore. Same as the first layer in our CNN architecture.
- Fully connected layer. Five hundred twelve neurons combined with ReLU layer, which is used to perform high-level reasoning like neural networks.

In the first convolutional layer, the input image patch is continuously filtered with 48 feature maps of  $3 \times 11 \times 11$  and  $3 \times 7 \times 7$  kernels with a stride of 2. The second convolutional layer continuously filters the output of the first convolutional layer with 128 feature maps of  $3 \times 9 \times 9$  and  $3 \times 5 \times 5$  kernels. The third convolutional layer filters the output of the second convolutional layer with 128 feature maps of  $3 \times 7 \times 7$  and  $3 \times 3 \times 3$  kernels, sequentially. The following fully connected layer has 512 neurons.

The softmax loss locates at the end of the CNN for the original image, which is expressed as:

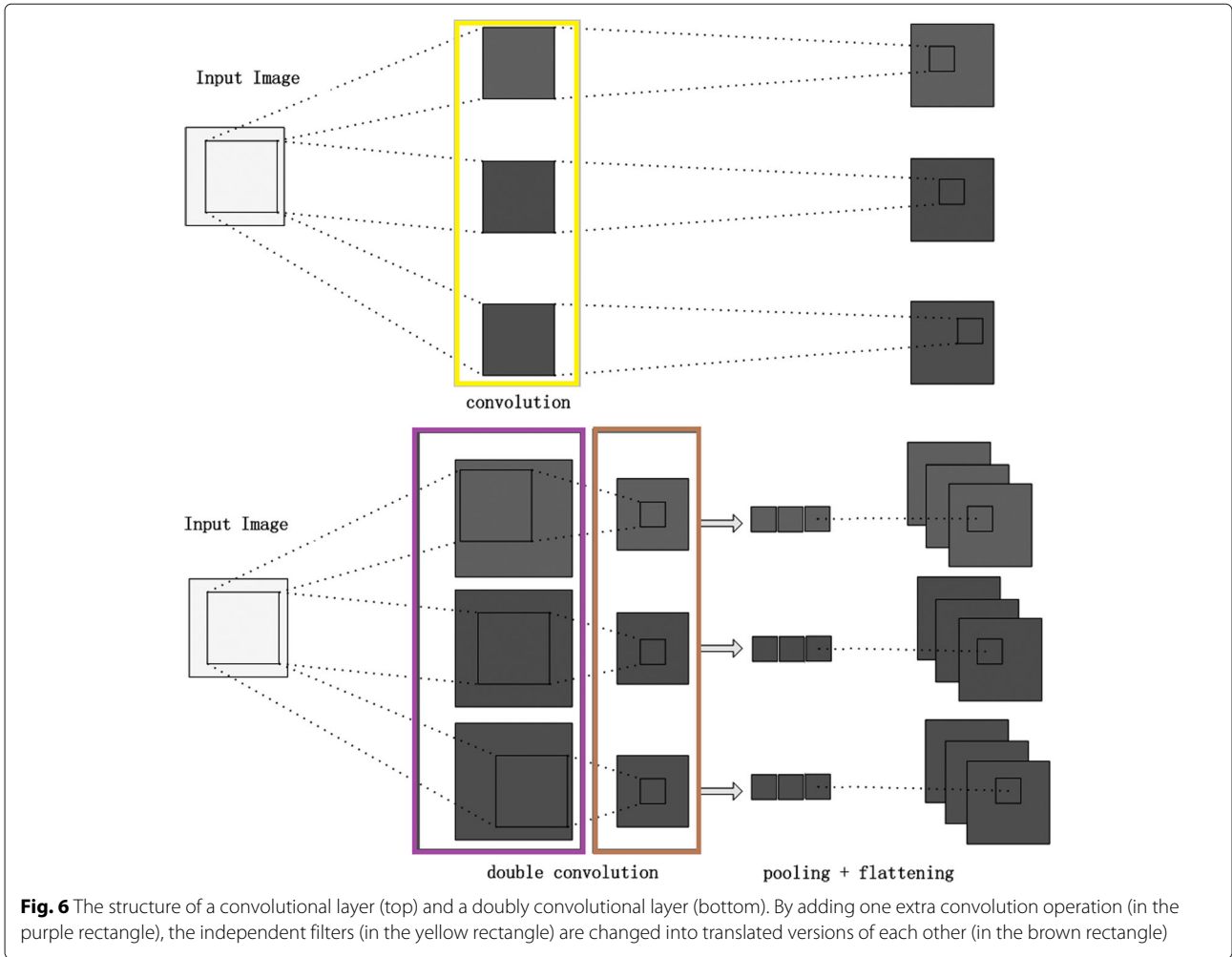
$$\text{SoftmaxLoss} = - \sum_{i=1}^N \log(P(\omega_k | L_i)), \quad (1)$$

where  $N$  denotes the number of the input images,  $P(\omega_k | L_i)$  indicates the probability of the  $k$ th image to image patch label  $L_i$  correctly classified as  $l_k$ .

Besides the softmax loss function used that has been exploited in other CNN, the triplet loss is also presented in our proposed CNN architecture.

$$\text{TripletLoss} = \frac{1}{2N} \sum_{i=1}^N \max(0, D(o_i, p_i) - D(o_i, n_i) + m), \quad (2)$$

where  $D(.,.)$  denotes the squared Euclidean distance between two  $l_2$  normalized vectors;  $o_i, p_i$ , and  $n_i$  respectively stand for the  $l_2$  normalized vectors from the original



**Fig. 6** The structure of a convolutional layer (top) and a doubly convolutional layer (bottom). By adding one extra convolution operation (in the purple rectangle), the independent filters (in the yellow rectangle) are changed into translated versions of each other (in the brown rectangle)

image, the positive image, and the negative image, as shown in Fig. 5;  $m$  denotes the hyperparameter to confine the value of Eq. (2) greater than zero as the difference between the original image and the positive image is expected to exceed the difference between the original image and the negative image.

The two losses mentioned above are then integrated into a weighted combination:

$$\text{Loss} = \lambda \text{SoftmaxLoss} + (1 - \lambda) \text{TripletLoss}, \quad (3)$$

where  $\lambda$  denotes the weight that is used to manipulate the trade-off between the softmax loss and the triplet loss.

The feature maps hierarchically extracted by our proposed CNN are shown in Fig. 7. The extracted features include but not limited to color, shape, and texture of the refrigerators. These extracted features consist of both local features and global features in the front-view images of the refrigerator.

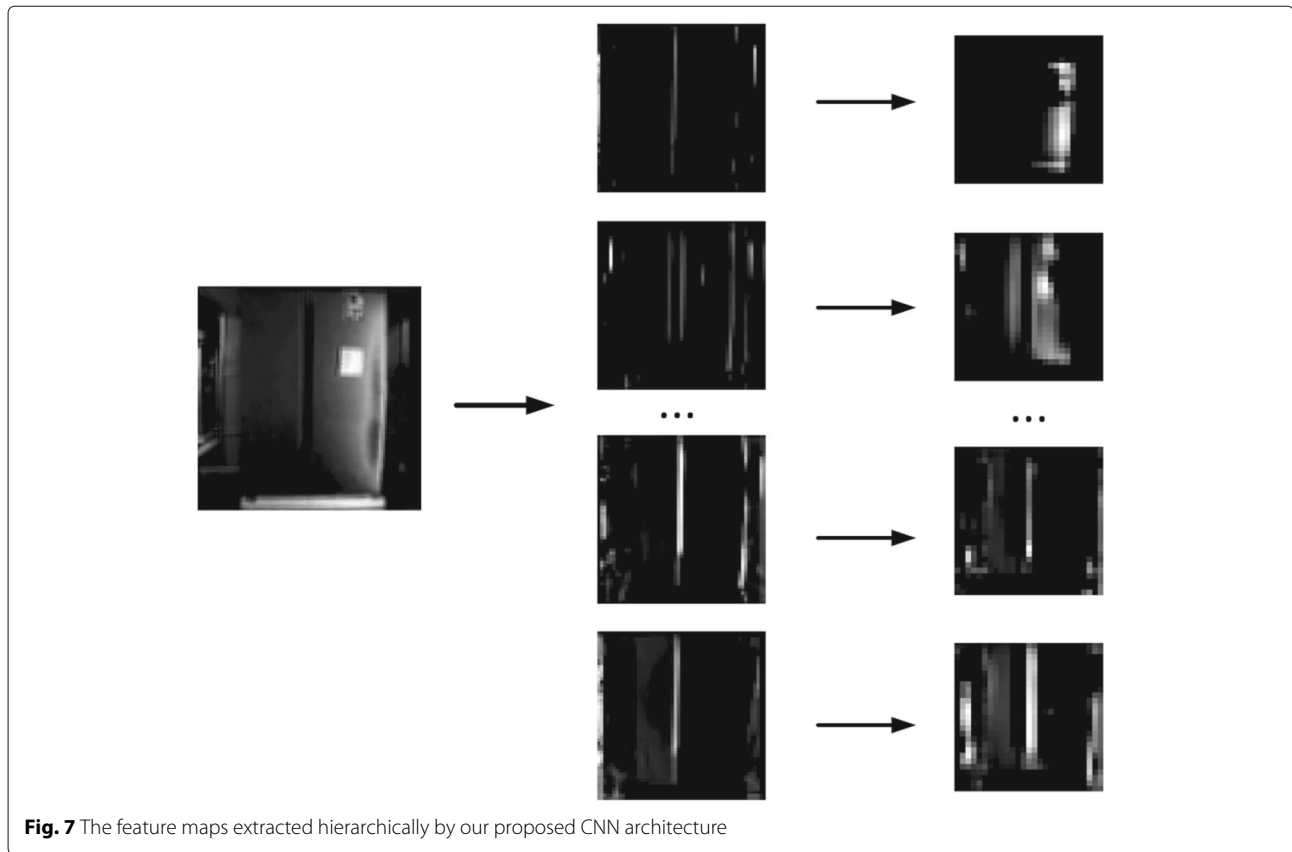
Figure 7 shows that in the first convolutional layer (the first column in Fig. 7), the global features including shape and edge of the refrigerator are extracted. In the

two following convolutional layers (the second and third columns in Fig. 7), local features are extracted hierarchically. Notable that these features, which are different from the handcrafted features exploited in [5–8], are automatically extracted through our proposed CNN.

### 3 Results and discussion

#### 3.1 Experimental samples and pre-processing

All of the refrigerator images were captured by using a digital camera (Canon EOS 760D, Japan); its setting is the manual mode, and the exposure level is 0.0 without zoom and flash. The color space used in the classification is the sRGB (standard RGB). The original size of the captured images is  $3200 \times 2400$  (0.1 mm/pixel), and they are stored into PNG (Portable Network Graphic) format. To reduce the over-fitting that may exist in a few images, we enlarge the dataset of the originally captured images of refrigerators with augmentation techniques, such as translations and vertical and horizontal reflections. Finally, there are totally 31,247 images of refrigerator belonging to 30 categories.



**Fig. 7** The feature maps extracted hierarchically by our proposed CNN architecture

After the procedure of data augmentation, the images are then resized into  $227 \times 227$ . Among all of the images, 20,000 of them are taken as training data, and the others are used as testing data.

### 3.2 Training

During the training process, each pair of the captured refrigerator image and its label of classification, a positive image and a negative image of the original image are taken as input to the CNN. The proposed CNN is updated with the back propagation mechanism, which calculates the softmax loss from the original image and the triplet loss from the triple input images. The training is performed on GPU of high performance; it takes  $10^5$  times of iterations. For each iteration, it costs around 2.4 s. Through training, our proposed CNN can establish the mapping between the input image of refrigerator and the output refrigerator category.

### 3.3 Experimental result

After the training process, we conducted experiments to testify the performance of our proposed CNN. The experimental result of the full testing data is shown in Fig. 8. From Fig. 8, we can obtain that the classification accuracy of our method reaches to 99.99% after about 3000 iterations. Meanwhile, the training loss of our method

decreases to 0.01. (The decreasing process of training loss represents the convergence of the proposed CNN architecture.)

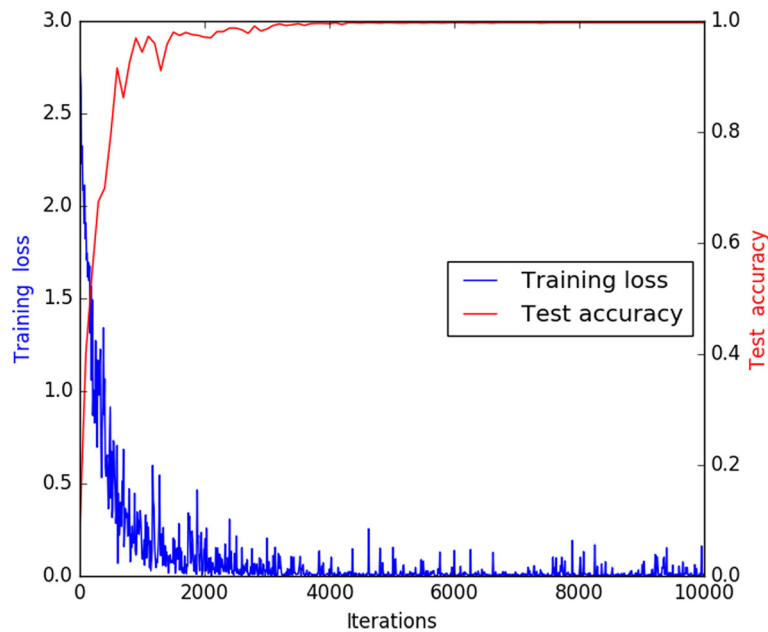
Notably, as numerous images of the refrigerator containing challenges including dense clutter, varying backgrounds, large homogeneous areas, subtle salient regions, and specular reflection are included in our training process, the negative effects brought by them to the image classification have been significantly eliminated by our method. Experimental results show that our method can handle these challenges well.

We conducted comparing experiments between the state-of-the-art SVM-based image classification methods [5–8], human's visual inspection, and our method. Different handcrafted features including single feature and composite features are exploited by these SVM-based methods, respectively.

Meanwhile, to quantitatively compare the performance of our method and the state-of-the-art SVM methods, the Precision and Recall are exploited, as expressed in Eqs. (4) and (5).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$



**Fig. 8** The experimental result of our proposed CNN architecture

where TP denotes true positive, FP is the false positive, TN is the true negative, and FN denotes the false negative. The comparing experimental results are presented in Table 1.

The second and third columns of Table 1 respectively shows the Precision and Recall of the full testing data. From these, we can conclude that our proposed method outperforms the state-of-the-art methods and human's visual system. Note that there are 30 categories of refrigerators in our experiments, which greatly confine the performance of human's visual inspection. While the number of classes is limited, the performance of human's visual system is promising.

To testify the performance of our proposed method in extreme situations, we intentionally capture images with artifacts as shown in Fig. 9.

The experimental results of these images are shown in the fourth and fifth columns of Table 1. These results show that the performance of our method on images with artifacts is better than the state-of-the-art methods, too.

There are misclassifications due to the severe artifacts in these images. Most of the misclassifications might be caused by extreme viewing conditions, e.g., the workers standing in front of the refrigerator. But experimental result shows that even under an extreme situation, our method can still obtain satisfactory outcomes.

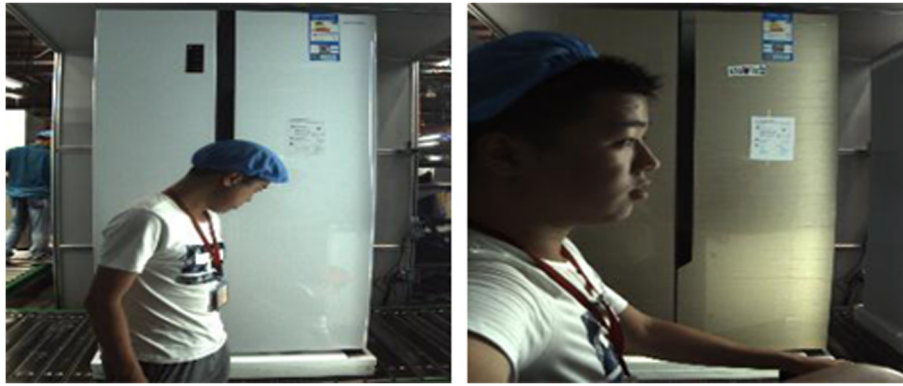
### 3.4 Analysis

As mentioned in Section 2.2, the local and global features are automatically extracted by our proposed CNN. Similar to human's visual system, our proposed CNN can extract the global features including the color, shape of the refrigerators and the local features from the display, crack, ice maker, and handle of the refrigerators. The global features integrated with local features form a layout for each category of refrigerators. The whole layout is utilized to establish the mapping between the input front-view image of the refrigerator and the output classification label.

**Table 1** The Precision and Recall of SVM methods, human's visual system, and our method

Methods	Full testing data		Partial data	
	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Gabor + SVM [5]	75.24	7.21	70.12	6.11
Wavelet + SVM [6]	76.48	6.89	73.37	5.93
Wavelet + Gabor + SVM [7]	78.19	6.37	77.65	5.31
Combined features + SVM [8]	80.21	6.01	80.33	5.75
Human's visual system	75.21	10.01	82.54	5.18
Our method	99.96	2.31	95.54	3.12





**Fig. 9** The testing data with different artifacts

Furthermore, different from the previous proposed CNN, our proposed CNN architecture with the introduced triplet loss significantly enhanced the accuracy of image classification by jointly optimizing the classification loss (between the label of the original image and the output classification result) and the similarity loss (between the original image, the positive image, and the negative image), as shown in Fig. 5. The parameter  $\lambda$  in Eq. (3) that is used to balance the softmax loss and triplet loss plays an essential role in the proposed CNN architecture. With  $\lambda$  is set to 1 or 0, the performance of the architecture would degenerate to softmax loss or triplet loss, respectively. According to the process of error and trial, it is reasonable to assign a value greater than 0.5 to  $\lambda$ , which shows that the softmax loss should be more important than the triplet loss in our proposed CNN architecture. However, the introduced triplet loss contributes to the image classification due to the complementary information from both the positive image and negative image. It is suitable for the characteristics of the front-view of refrigerator as dense clutter, varying backgrounds, subtle salient regions, and the specular reflection in the images of refrigerators all can be extracted and represented by the combination of the softmax loss and triplet loss, and the subtle difference between intra-category and inter-categories of refrigerators can be differentiated from each other. Meanwhile, the doubly convolutional layer presented in our approach also positively affects the outcome of the proposed CNN architecture. It not only enhances the parameter-sharing efficiency but also accommodates the characteristics of the captured images. Because the refrigerators were transported on the same conveyor belt and the position of the camera was fixed, the captured intra-class refrigerators are translated versions of each other, which is suitable for the feature maps learned from doubly convolutional layers.

The compared experimental results in Fig. 8 shows that our proposed method outperforms the state-of-the-art

image classification methods. It achieves better accuracy of image classification of refrigerators. Due to the accuracy of classification by our method, it can satisfy the practical requirement of industrial automation of refrigerator.

#### 4 Conclusions

In this paper, we propose a CNN-based image classification method based on the front-view images of the refrigerator to cope with the present challenges including dense clutter, varying background, large homogeneous areas, and specular reflection. Our experimental results validate the ability of the proposed technique in solving the refrigerator classification task in practical industrial automation.

Meanwhile, we have the following significant contributions. First of all, this is the first time to introduce machine-learning-based method into the classification of refrigerator instead of human's visual system. Secondly, the triplet loss and doubly convolutional layer both introduced by our method positively affect the accuracy of image classification. To our best knowledge, this is the early application of both the triplet and double convolution operation into CNN-based classifier. Thirdly, similar to human's visual system, our proposed CNN can extract the multi-scale features including both the local and global features of the front-view-based images of the refrigerator. As the extraction of multi-scale features of the image plays a vital role in human's visual system, our method can implement it automatically. Finally, our approach performs with an impressive superiority to both the state-of-the-art image classification techniques and human's visual inspection. Our proposed CNN-based method is a potential tool for image classification of the refrigerator and other images with similar characteristics as the front-view of the refrigerator.

In the future, we will continue to explore the implicit process of feature extraction by CNN and attempt to

find the explicit mapping between the performance of the CNN and the role of local and global features extracted by CNN. Meanwhile, we will also research on more applications of our proposed CNN-based classification method in various fields including medical image processing, object classification, and recognition [25–29].

#### Abbreviations

CNN: Convolutional neural network; SVM: Support vector machine; ReLU: Rectified linear unit; sRGB: Standard RGB; PNG: Portable Network Graphic; TP: True positive; FN: False negative; FP: False positive; TN: True negative

#### Acknowledgements

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

#### Funding

This research is supported, in part, by the National Natural Science Foundation of China under Grant Nos. 61572295, 61573212, and 61272241; the Natural Science Foundation of Shandong Province of China under Grant Nos. ZR2014FM031 and ZR2013FQ014; the Shandong Province Independent Innovation Major Special Project No. 2015ZDXX0201B03; the Shandong Province Key Research and Development Plan Nos. 2015GGX101015 and 2015GGX101007; and the Fundamental Research Funds of Shandong University No. 2015JC031.

#### Availability of data and materials

We can provide the data.

#### Authors' contributions

All authors took part in the discussion of the work described in this paper. BG put forward the main idea. YG wrote the first version of the paper and did part experiments of the paper. JL and BG revised the paper in different version of the paper. The contributions of the proposed work are mainly in two aspects: To our best knowledge, our work is the first one to implement the refrigerator classification with high accuracy. (1) In this paper, we propose a novel CNN-based classification framework designed specifically for classifying the refrigerators with defects in the input images. (2) Experimental results show that our system could improve the perform in accuracy for the state-of-the-art methods. All authors read and approved the final manuscript.

#### Authors' information

1. Yongchang Gao. He is now a doctor candidate in Shandong University. His research interests include information management, machine vision, and deep learning.
2. Jian Lian. He is now an instructor in Shandong University of Science and Technology. His interests include machine learning and image processing.
3. Bin Gong. He is now a professor in Shandong University. His interests include machine learning and image processing.

#### Ethics approval and consent to participate

Approved.

#### Consent for publication

Approved.

#### Competing interests

The authors declare that they have no competing interests. And all authors have seen the manuscript and approved to submit to your journal. We confirm that the content of the manuscript has not been published or submitted for publication elsewhere.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 July 2018 Accepted: 5 September 2018

Published online: 24 September 2018

#### References

1. C. S. Press, *National Bureau of Statistics of China, China statistical yearbook 2016*. (China Statistics Press, Beijing, 2016)
2. H. Akbar, N. Suryana, F. Akbar, Surface defect detection and classification based on statistical filter and decision tree. *Int. J. Comput. Theory Eng.* **5**(5), 774 (2013)
3. J. M. Lillocastellano, I. Morajimenez, C. Figuerapozuelo, J. L. Rojoalvarez, Traffic sign segmentation and classification using statistical learning methods. *Neurocomputing*. **153**, 286–299 (2015)
4. G. M. Farinella, D. Allegra, M. Moltisanti, F. Stanco, S. Battiato, Retrieval and classification of food images. *Comput. Biol. Med.* **77**, 23–39 (2016). <https://ieeexplore.ieee.org/document/4409066/>
5. Z. Sun, G. Bebis, R. Miller, in *International Conference on Digital Signal Processing*. On-road vehicle detection using gabor filters and support vector machines (IEEE, Los Alamitos, 2002), pp. 1019–1022
6. Z. Sun, G. Bebis, Miller R., in *International Conference on Control, Automation, Robotics and Vision. IEEE*. Quantized wavelet features and support vector machines for on-road vehicle detection, vol. 3, (2002), pp. 1641–1646
7. Z. Sun, G. Bebis, R. Miller, in *The IEEE, International Conference on Intelligent Transportation Systems, 2002. Proceedings. IEEE*. Improving the performance of on-road vehicle detection by combining gabor and wavelet features, (2002), pp. 130–135
8. X. Wen, L. Shao, W. Fang, Y. Xue, Efficient feature selection and classification for vehicle detection. *IEEE Trans. Circ. Syst. Video Technol.* **25**(3), 508–517 (2015)
9. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition. *Proc. IEEE*. **86**(11), 2278–2324 (1998)
10. G. E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets. *Neural Comput.* **18**(7), 1527–1554 (2006)
11. Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle, et al., Greedy layer-wise training of deep networks. *Adv. Neural Inf. Process. Syst.* **19**, 153 (2007)
12. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. *Int. Conf. Learn. Represent.* (2015)
13. Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, in *Proceedings of the ACM International Conference on Multimedia, ACM*. Caffe: Convolutional architecture for fast feature embedding (ACM, New York, 2014), pp. 675–678
14. S. C. Turaga, J. F. Murray, V. Jain, F. Roth, M. Helmstaedter, K. Briggman, W. Denk, H. S. Seung, Convolutional networks can learn to generate affinity graphs for image segmentation. *Neural Comput.* **22**(2), 511–538 (2010)
15. A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, M. Nielsen, in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network (Springer, Berlin, 2013), pp. 246–253
16. J. Long, E. Shelhamer, T. Darrell, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Fully convolutional networks for semantic segmentation (IEEE, Los Alamitos, 2015), pp. 3431–3440
17. A. Krizhevsky, I. Sutskever, G. E. Hinton, in *Advances in neural information processing systems*. Imagenet classification with deep convolutional neural networks (MIT Press, Cambridge, 2012), pp. 1097–1105
18. R. Socher, B. Huval, B. Bath, C. D. Manning, A. Y. Ng, in *Advances in Neural Information Processing Systems*. Convolutional-recursive deep learning for 3D object classification (MIT Press, Cambridge, 2012), pp. 665–673
19. K. Simonyan, A. Vedaldi, A. Zisserman, Deep inside convolutional networks: visualising image classification models and saliency maps. arXiv preprint arXiv:1312.6034. <http://cn.arxiv.org/abs/1312.6034>
20. A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, L. Fei-Fei, in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. Large-scale video classification with convolutional neural networks (IEEE, Los Alamitos, 2014), pp. 1725–1732
21. W. Hu, Y. Huang, L. Wei, F. Zhang, H. Li, Deep convolutional neural networks for hyperspectral image classification. *J. Sensors*. **2015**, 1–12 (2015)
22. E. Ahn, A. Kumar, J. Kim, et al, in *IEEE, International Symposium on Biomedical Imaging. IEEE*. X-ray image classification using domain transferred convolutional neural networks and local sparse spatial pyramid, (2016), pp. 855–858
23. J. Shi, S. Zhou, X. Liu, et al., Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset. *Neurocomputing*. **194**(C), 87–94 (2016)

24. Z. Liu, P. Luo, X. Wang, et al., Deep learning face attributes in the wild// IEEE International Conference on Computer Vision. IEEE Comput. Soc, 3730–3738 (2015)
25. Y. Zhang, J. Lian, M. Fan, et al., Deep indicator for fine-grained classification of banana's ripening stages. *Eurasip J. Image Video Process.* **2018**(1), 46 (2018)
26. J. Lian, Y. Zheng, W. Jiao, et al, Deblurring sequential ocular images from multi-spectral imaging (MSI) via mutual information. *Med. Biol. Eng. Comput.* **56**(6) (1107)
27. J. Lian, Y. Zheng, P. Duan, et al., Measuring spectral inconsistency of multispectral images for detection and segmentation of retinal degenerative changes. *Sci Rep.* **7**(1), 11288 (2017)
28. H. Pu, M. Fan, Yang J., et al, Quick response barcode deblurring via doubly convolutional neural network. *Multimedia Tools Appl* (2018). (2) 1–6
29. J. Lian, S. Hou, X. Sui, et al., Deblurring retinal optical coherence tomography via convolutional neural network with anisotropic and double convolution layer. *Iet Comput. Vis.* (2018)

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---