Orientation Detection Using a CNN Designed by Transfer Learning of AlexNet

Fusaomi Nagata[a,*], Kohei Miki[a], Yudai Imahashi[a], Kento Nakashima[a]
Kenta Tokuno[a], Akimasa Otsuka[a], Keigo Watanabe[b] and Maki K. Habib[c]

[a]Department of Mechanical Engineering, Faculty of Engineering, Sanyo-Onoda City University
Sanyo-Onoda 756-0884, Japan
[b]Graduate School of Natural Science and Technology, Okayama University
Okayama 700-8530, Japan
[c]Mechanical Engineering Department, School of Sciences and Engineering, The American University in Cairo
AUC Avenue, P.O.Box 74, New Cairo 11835, Egypt
*Corresponding Author: nagata@rs.socu.ac.jp

Abstract

Artificial neural network (ANN) which has four or more layers structure is called deep NN (DNN) and is recognized as a promising machine learning technique. Convolutional neural network (CNN) has the most used and powerful structure for image recognition. It is also known that support vector machine (SVM) has a superior ability for binary classification in spite of only two layers. We have developed a CNN&SVM design and training tool for defect detection of resin molded articles, and the effectiveness and validity have been proved through several CNNs design, training and evaluation. The tool further enables to easily design a CNN model based on transfer learning concept. In this paper, a CNN acquired by transfer learning of AlexNet, which is the winner of ImageNet LSVRC2012, is designed to recognize the orientation of objects. The effectiveness of the transfer learning based CNN is evaluated using test data set.

Keywords: Deep learning, Convolutional neural network, Defect detection, Transfer learning, Orientation detection, robotic pick and place.

1. Introduction

Artificial neural network (ANN) which has four or more layers structure is called deep NN (DNN) and is recognized as a promising machine learning technique. Convolutional neural network (CNN) has the most used and powerful structure for image recognition. It is also known that support vector machine (SVM) has a superior ability for binary classification in spite of only two layers. Nagi et al. designed max-pooling convolutional neural networks (MPCNN) for vision-based hand gesture recognition[1]. The MPCNN could classify six kinds of gestures with 96% accuracy and allow mobile robots to perform real-time gesture recognition. Weimer et al. also designed a deep CNN architectures for automated feature extraction in industrial inspection process[2]. The CNN automatically generates features from massive amount of training image data and demonstrates excellent defect detection results with low false alarm rates. Faghih-Roohi et al. presented a different type of deep CNN for automatic detection of rail surface defects[3]. It was concluded that the large CNN model performed a better classification result than the small and medium CNN, although the training required a longer time. Zhou et al. used a CNN to classify the surface defects of steel sheets[4]. The CNN could directly learn better representative features from labeled images of surface defects. Further, Ferguson et al. presented a system to identify casting defects in X-ray images based on the Mask Region-based CNN architecture[5,6]. It is reported that the proposed system simultaneously performed defect detection and segmentation on input images making it suitable for a range of defect detection tasks.

We have developed a CNN&SVM design and training tool for defect detection of resin molded articles and the effectiveness and validity have been proved through several CNNs design, training and evaluation[7,8,9]. The tool further enables to easily design a CNN model based on transfer learning concept. When industrial robots are applied to pick and place tasks of resin molded articles, information of each

object's position and orientation is essential. Recognition and extraction of object position in an image are not so difficult if using image processing technique, however, that of orientation is not easy due to the variety in shape. In this paper, a CNN acquired by transfer learning of AlexNet, which is the winner of ImageNet LSVRC2012, is introduced to recognize the orientation of objects in images. The effectiveness of the CNN is evaluated using test image data set of thin resin mold articles.

2. Design Tool for CNN and SVM

Figure 1 shows the main dialogue of the developed CNN&SVM design tool. In training of CNN, pre-training using randomly initialize weights and additional (successive) training with once trained weights can be selected. As for SVM, one-class unsupervised learning and two class supervised learning can be selectively executed. Also, favorite CNN, which is used for a feature extractor, and Kernel function are selected.

The tool has another promising function to design original CNNs based on transfer learning. For example, the following main items can be set for the operation of transfer learning through the dialogue.

(1) Folders for training and test images.

(2) Base CNNs used for transfer learning such as AlexNet, VGG16, VGG19, GoogleNet and Inception-V3.

(3) Learning parameters such as max epochs, mini batch size, desired accuracy and loss, learning rates for convolution layers and fully connected layers.

The software shown in Fig. 1 is developed on MATLAB system optionally installed with Neural Network Toolbox, Parallel Computing Toolbox for GPU, Deep Learning Toolbox, Statistics and Machine Learning Toolbox.

3. Images for Training and Test

Training image generator was already proposed to efficiently augment limited number of training images[7]. By using the generator, images for training are prepared considering typical twelve orientations, i.e., 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150 and 165 degrees. Figures 2 and 3 show examples of the training images for the categories of 45 and 165 [deg], respectively. The resolution and channel are 200×200 and 1, respectively.
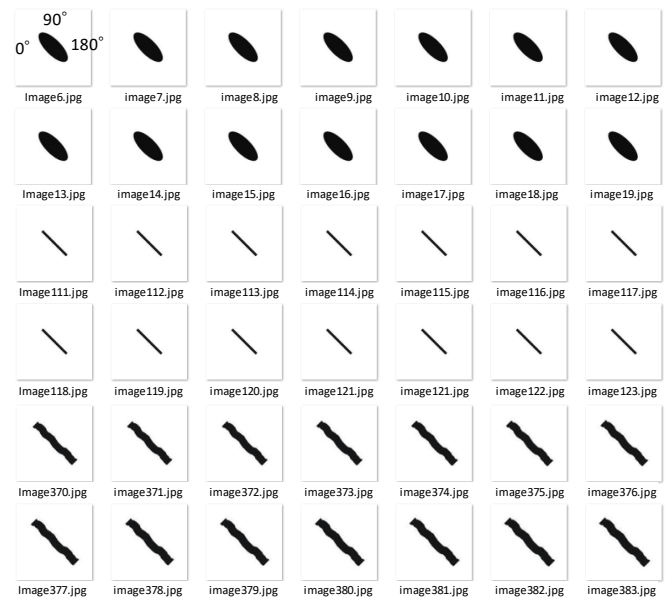

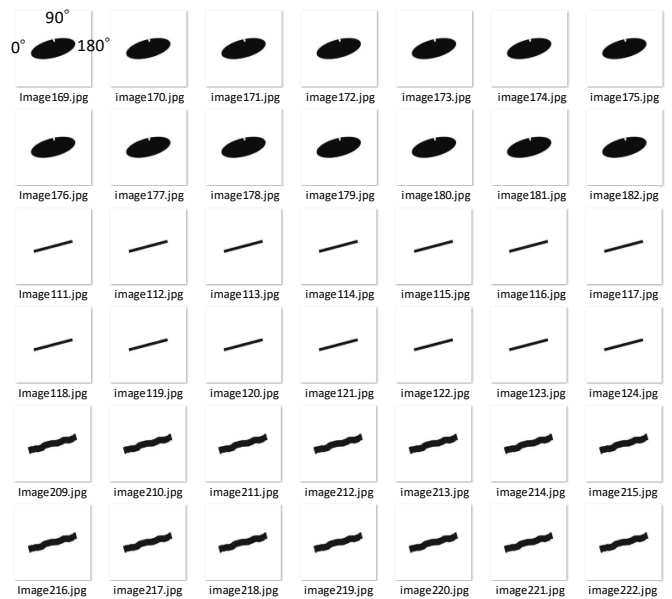Fig. 2. Example of training images for 45 [deg].

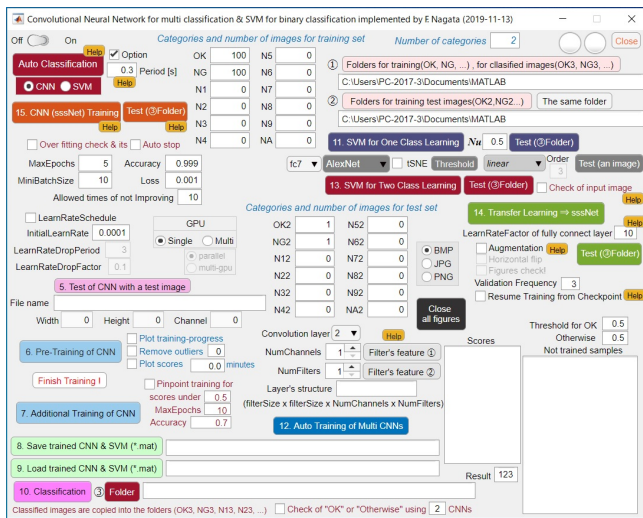Fig. 3. Example of training images for 165 [deg].

Fig. 1. Main dailogue developed for CNN and SVM design and training.
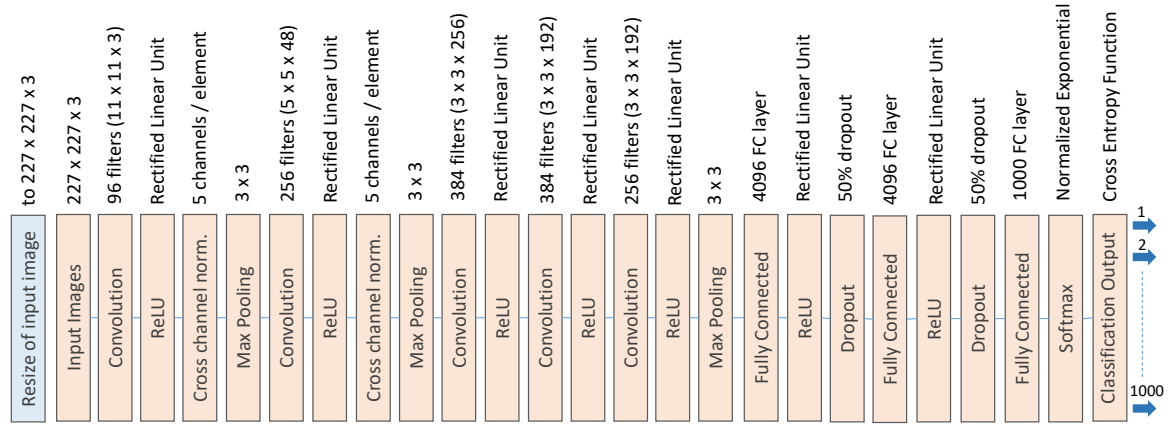
Fig. 4. Network structure of well-known CNN named AlexNet.

4. Transfer Learning Based CNN

In this section, a transfer learning based CNN is designed to learn the feature of orientation included in images as shown in Figs. 2 and 3. Figure 4 illustrates the structure of the original AleXnet. In order to make the CNN have an ability to classify input images into 12 categories as 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150 and 165 [deg], the fully connected layers are replaced as shown in Fig. 5 before executing transfer learning. 6,889 images consisting of 12 categories are used for the transfer learning. As for training parameters, mini batch size is given 50. Iteration is the number of mini batches needed to complete one epoch, so that one epoch in this transfer learning is composed of 6,889/50≒137 iterations. Desired accuracy and loss are set to 1 and 0, respectively. Besides, learning rates of convolutional layers and fully connected ones are set to 0.0001 and 0.001, respectively. It is important for fast and stable convergence in transfer learning to set the learning rate in convolutional layers smaller than that of fully connected layers.

The training was conducted using a single PC with a Core i7 CPU and a GPU (NVIDIA GeForce GTX 1060, 6GB). The training progress is shown in Fig. 6, in which both the training accuracy and loss seem to well converge to desired values. It actually took about 40 minutes until the learning was stopped since both the accuracy and loss had not been improved during 10 consecutive iterations or more. Note that this training could be completed within one epoch by severally giving different learning rates in convolutional layers and fully connected layers. Through the process

explained above, an original CNN model acquired by transfer learning of AlexNet, which is the winner of ImageNet LSVRC2012, is presented to recognize the
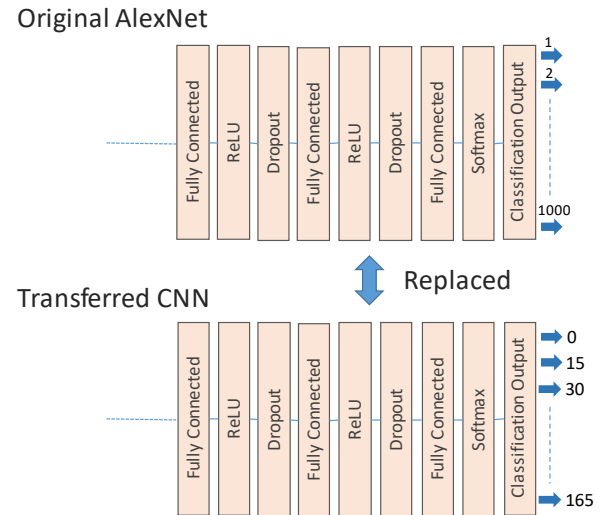

Fig. 5. Replacement of fully connected layers for dealing with target classification task, i.e., 12 categories.
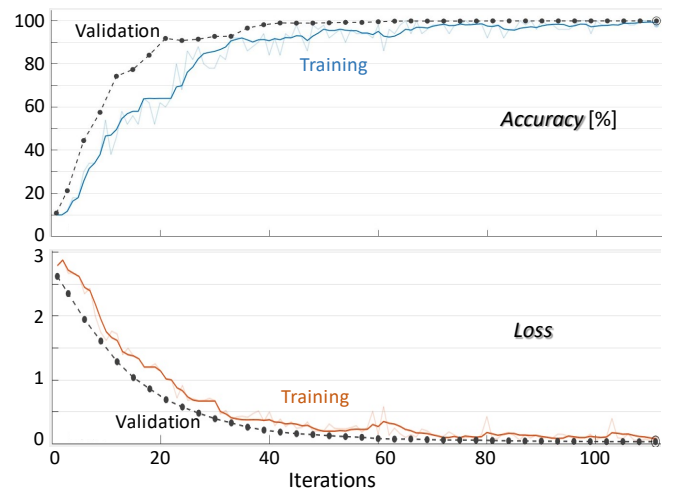

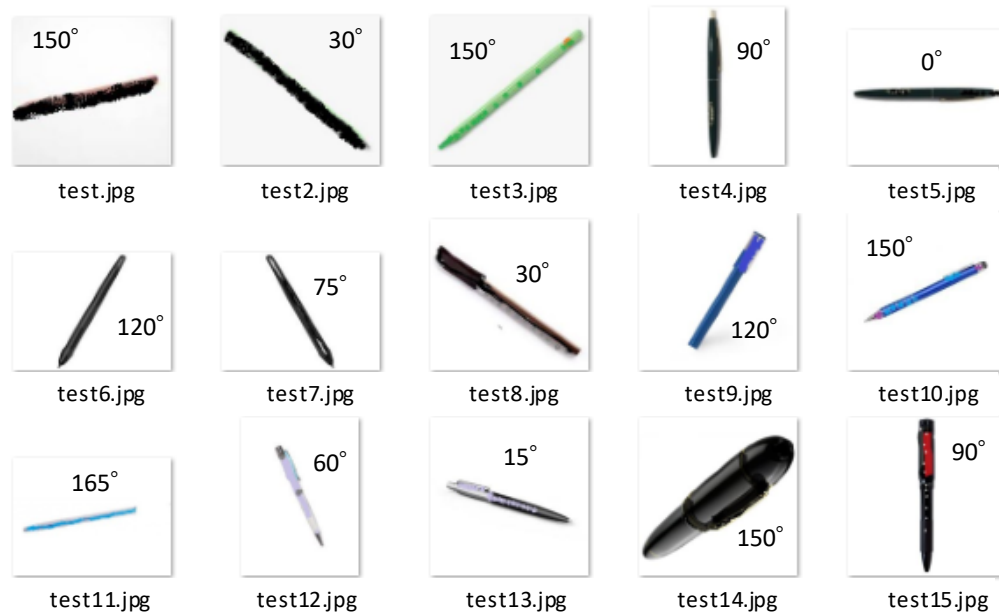Fig. 6. Training progress of transfer learning shown in Fig. 5.

Fig. 7. Classification results using transfer learning based CNN.

orientation of objects.

After the training, the generalization ability of the transfer learning based CNN is checked using 15 test images imitating resin molded articles which have not been included in the training data set. Figure 7 shows the photos and their classification results, i.e., the angle shown in the JPEG images is the output from the CNN. It is observed from the results that the obtained CNN has a promising ability that can recognize the orientations of objects in the images. However, some visual inconsistencies, e.g., between 'test7.jpg' (75°) and 'test12.jpg' (60°); 'test.jpg' (150°) and 'test3.jpg' (150°) are observed. As can be clearly seen, some images in Fig. 7 are not square. That is the reason why the main cause of these results seems to be the conversion of resolution before classification. The resolution of images given to the input layer has to be converted to $227 \times 227 \times 3$ fixed according to the input layer of the AlexNet, which brings out some undesirable deformation of images and the resultant ambiguities in classification.

5. Conclusions

In this paper, a CNN acquired by transfer learning of AlexNet, which is the winner of ImageNet LSVRC2012, is introduced to recognize the orientations of objects. The effectiveness and promise of the CNN are evaluated using test images imitating thin resin mold articles. In future work, we are planning to apply a small robot incorporated with this CNN to an actual production line of resin molded articles with various shapes. Because orientation information is essential for the robot to successfully play a pick and place task.

References

(1) J. Nagi, F. Ducatelle, G.A.D. Caro, D. Ciresan, U. Meier, A. Giusti, F. Nagi, J. Schmidhuber, and L.M. Gambardella : "Max-pooling convolutional neural networks for vision-based hand gesture recognition," 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA2011), pp. 342–347, 2011

(2) D. Weimer, B. Scholz-Reiter, and M. Shpitalni : "Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection," CIRP Annals – Manufacturing Technology, Vol. 65, No. 1, pp. 417–420, 2016

(3) S. Faghih-Roohi, S. Hajizadeh, A. Nunez, R. Babuska, and B.D. Schutter : "Deep convolutional neural networks for detection of rail surface defects," Procs. of the 2016 International Joint Conference on Neural Networks (IJCNN2016), Vancouver, Canada, pp. 2584–2589, July 2016

(4) S. Zhou, Y. Chen, D. Zhang, J. Xie, and Y. Zhou : "Classification of surface defects on steel sheet using convolutional neural networks," Materials and Technology, Vol. 51, No. 1, pp. 123–131, 2017

(5) M. Ferguson, R. Ak, Y. Lee, and K. Law: "Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning," Smart and Sustainable Manufacturing Systems, Vol. 2, No. 1, pp. 137–164, 2018

(6) K. He, G. Gkioxari, P. Dollr, and R. Girshick, "Mask R-CNN," Procs. of 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2980–2988, 2017

(7) F. Nagata, K. Tokuno, H. Tamano, H. Nakamura, M. Tamura, K. Kato, A. Otsuka, T. Ikeda, K. Watanabe, M.K. Habib, "Basic application of deep convolutional neural network to visual inspection," Procs. of International Conference on Industrial Application Engineering (ICIAE2018), pp. 4–8, Okinawa, March 2018

(8) F. Nagata, K. Tokuno, K. Nakashima, A. Otsuka, T. Ikeda, H. Ochi, K. Watanabe, M. K. Habib, "Fusion method of convolutional neural network and support vector machine for high accuracy anomaly detection," Procs. of the 2019 IEEE International Conference on Mechatronics and Automation (ICMA 2019), pp. 970–975, Tianjin, China, August 2019

(9) F. Nagata, K. Tokuno, K. Mitarai, A. Otsuka, T. Ikeda, H. Ochi, K. Watanabe, M. K. Habib, "Defect detection method using deep convolutional neural network, support vector machine and template matching techniques," Artificial Life and Robotics, 8 pages, First Online: 02 July 2019