Deep Learning Model for Determining Defects of Vision Inspection Machine Using Only a Few Samples

Bo-lin Jian,[1] Jui-Pin Hung,[2] Cheng-Chi Wang,[2*] and Chun-Chang Liu[2]

[1]Department of Electrical Engineering, National Chin-Yi University of Technology, Taichung 41170, Taiwan
[2]Graduate Institute of Precision Manufacturing, National Chin-Yi University of Technology, Taichung 41170, Taiwan

For the intelligent manufacturing field, after finishing the cutting process, a metal surface may have various defects such as scratches, residues, and dirt. However, the conventional method of determining defects has the disadvantages of being time-consuming and expensive. In addition, it is necessary to consider the cost of collecting samples and the labor cost when practically collecting samples from industries. Therefore, in this study, we optimized the determination of the defects of the production component by a deep learning (DL) model with a few samples and used an image sensor to take pictures of the specific area of the component. Meanwhile, an entropy calculation method is proposed to determine the most suitable kernel size of a convolution layer. We analyzed and established a deep learning model to determine whether the finished products of a vision inspection machine have defects using only a few samples. We compared the pros and cons of DarkNet-53, which is a convolutional neural network (CNN) that is 53 layers deep, and AlexNet, which is a deep CNN, with the DenseNet-201 model in the experiments. The obtained experimental results indicate that the proposed method can effectively increase the rate of recognition between defective and nondefective samples and reduce the training cost. The results of this paper may contribute to the development of a novel diagnostic technique and also be helpful for the intelligent manufacturing industry.

1. Introduction

Nowadays, manufacturers are devoting much of their effort to enhancing product quality and production efficiency by especially visual inspection before shipment. If this type of inspection is carried out by manual measurement, it will not only require considerable labor, but it may also result in inaccurate testing results and affect the quality of the shipped products owing to human factors such as fatigue and manual measurement errors. Therefore, automated optical inspection has started to be increasingly used. Nowadays, deep learning (DL) is increasingly used as it overcomes hardware limitations. In particular, conventional image processing

combined with DL detection is becoming increasingly popular. Integrating a DL model with existing optical inspection systems is a technical breakthrough relieving a bottleneck of defect inspection systems in the manufacturing industry, and it also achieves the objective of advancing manufacturing processes. Given the trend of smart manufacturing, how to quickly adjust to new and complex manufacturing processes has become a very important subject. In studies of this subject, industrial machine vision inspection and DL are combined. Results of such studies are expected to lead to the development of a visual inspection technology with high adaptability to overcome the bottleneck of current image processing technologies and to the advancement of manufacturing processes.

The modern image classification technology in DL provides a new solution for automated image inspection and it can enhance the accuracy of image determination.[1,2] In addition, the application of artificial intelligence (AI) has become possible with the significant advances of graphic processing unit (GPU) hardware. Some renowned DL frameworks, such as DarkNet,[3] TensorFlow,[4] and PyTorch,[5] have also been developed, and they have facilitated technological advancements in the DL domain. Many studies adopt convolutional neural network (CNN) models to automatically classify different images.[6] On the basis of the DL structure, CNN models are easy to train and can automatically search for useful features.[7,8] DL technology has been used to inspect and determine various types of image defect, and it has been proven to be a very effective approach.[9]

In addition to discussing the effectiveness of automated feature extractors in DL, their efficiency must also be considered. The well-known feature extractors at present include Resnet-101, Resnet-152, DarkNet-19, and DarkNet-53, which are deep residual networks.[10] The model structure of DarkNet-19[11] is similar to that of the Visual Geometry Group (VGG).[12] DarkNet-19 is a CNN with 19 convolutional layers and 5 maxpooling layers. DarkNet-53 mixes elements of DarkNet-19 and Resnet.[10] DarkNet-53[13] is much more effective than DarkNet-19,[11] and its efficiency is about 1.5 times higher than that of Resnet-101.[14] In cases where DarkNet-53 and Resnet-152[10] have the same effect, DarkNet-53 has efficiency that is twice as high as that of Resnet-152.[15] At the same time, DarkNet-53 achieves the highest calculation speed in terms of floating points per second among the network structures, which indicates that its network structure can make the best use of GPU.[13] Deep networks have convergence difficulty. Many researchers have changed the activation function in order to prevent gradient disappearance, but issues still remain.[16,17] For example, activation functions such as Type-2 Fuzzy[18] and ReLTanh[19] have been adopted. However, the disappearance or explosion of the gradient is probably due to the high nonlinearity of the deep network. DarkNet-53, Resnet-101, and Resnet-152 use the residual learning method to solve the problem of accuracy that increases first and then saturates, but this also results in redundant networks.[20] Therefore, in this study, we focus on the determination of the defects of the product using the DL model with a few samples and also use an image sensor (ex: camera) to take pictures of a specific area of the product and propose a parameter adjustment method using only a few samples. This method determines the minimum number of network layers used and the minimum kernel size of a CNN to solve the problems of redundant network layers and selecting the optimal kernel size to enhance the recognition rate of products and shorten the model

training time. It is hoped to upgrade the diagnosis technology for intelligent manufacturing industry and decrease the cost of production.

2.   Methods

The research process proposed in this research is shown in Fig. 1. The experimental setup is introduced in Sect. 2.1, which includes the description of hardware and inspected objects for the machines, the specifications and photo shooting angles of the imaging equipment, the training computer of DL, and the specifications for the implemented computers. In Sect. 2.2, the number of images and image preprocessing methods used by the DL model are described. In Sect. 2.3, the principles of DarkNet-53 and the procedures for improving the model are described.
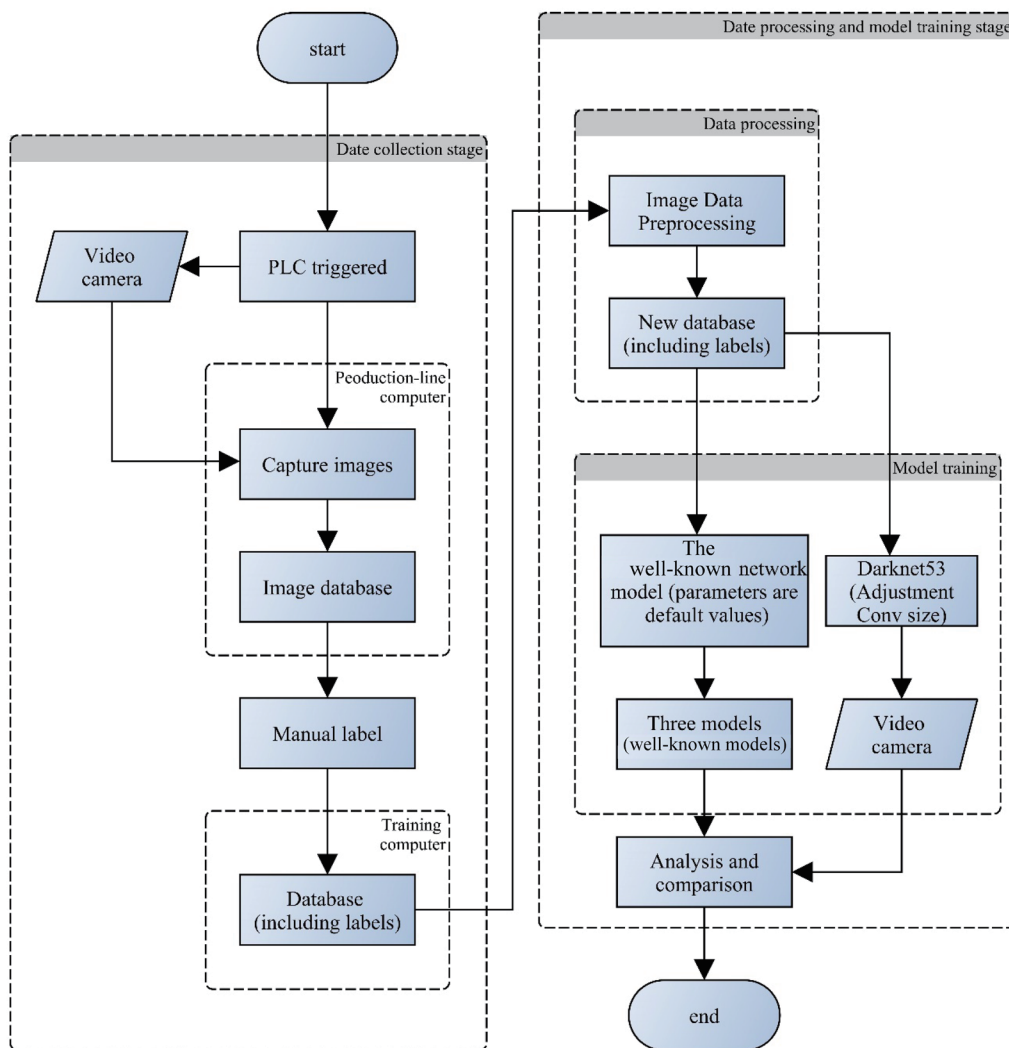

Fig. 1.   (Color online) Flowchart of experiment.

2.1   Experimental setup

2.1.1   Industrial automation equipment

Firstly, we discuss the hardware structure of the 506 Vision Inspection Machine. A photo of the machine is shown in Fig. 2. The machine includes 13 inspection and backup stations (reserved for use), an automatic detection feeding station, a transfer station (transfer components from station to station using the turntable), and the station for separating normal and defective materials. In this research, we focus on one of the distinctive stations named the small end-face vision inspection station. The appearance of the inspected object and its position detected by the small end-face vision inspection station are shown in Fig. 3. This figure also shows a diagram of the mechanism of object inspection. As this component is viewed from the direction of the smaller diameter end (area enclosed in red dashed line), it is named the small end face. The red plane of the component in the figure is the area inspected by the small end-face vision inspection station.

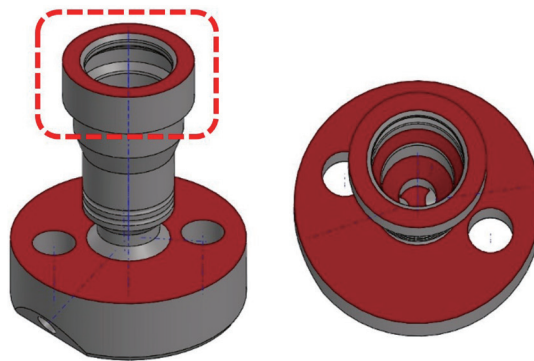Fig. 2.   (Color online) 506 Vision Inspection Machine.

Fig. 3.   (Color online) The area enclosed in a red dashed line is the small end-face portion of component 506.

2.1.2   Image-capturing device

The small end-face vision inspection station is used to determine the defects in the two planes and inside the holes of the small end face of component 506, and it takes multiple images at different light angles and focal distances. The small end-face vision inspection station includes mechanical moving parts and machine vision inspection equipment parts. The movement control of mechanical parts and two CCD cameras are integrated with the serial port built-in CPU module (KV-3000; KEYENCE) and programmable logic controller (PLC); the vision inspection machine structure is shown in Fig. 4.

The vision inspection machine structure has the following mechanical movement components: two-step motors, a coupling device, a circulation system, a circulation fixture base, a guideway, a reed switch, and a fixture pallet. This structure controls the movement of the components and the light source through the collaboration of the motors and circulation system. Each component is irradiated by the ring light at three different angles, and the height of the component is adjusted after taking the photos of the top surface of the component at three different light source angles, and the focal distance of the top CCD camera (CCD1) is then adjusted with respect to the lower surface portion of the component. At this moment, the ring light is adjusted again to take photos of the lower surface portion of the component at three different light source angles. At the same time, the side CCD camera (CCD2) also takes separate photos of the side of the hole. After all the images are taken, the component is moved back to the fixture pallet and allowed to standby for subsequent operations.
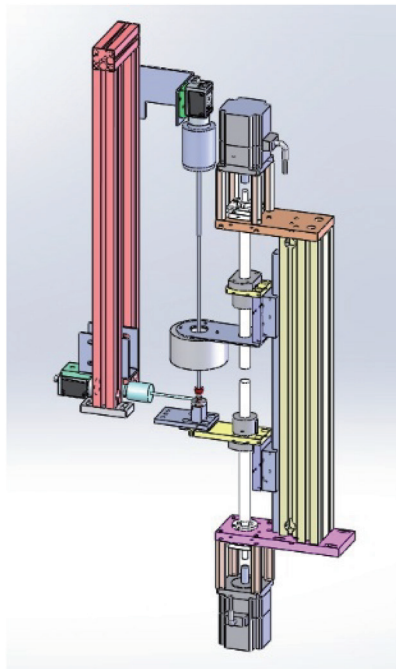

Fig. 4.   (Color online) Vision inspection machine structure.

For the vision inspection machine, we choose low-noise, high-sensitivity CCD cameras for imaging, and we use the standard interface GigE of industrial cameras for high-speed video transmission. We control the PLC transmission via Ethernet. We detect defects in this study by observing collision traces or scratches on the plane; thereby, we choose a ring light source that can highlight the plane features, and the types of defect are distinguished by taking photos at different light source angles. The brightness of the light source is controlled by the light source controller. The specifications of the vision inspection machine components are shown in Table 1. The schematic diagram for the height adjustment of the light source and the movement of the CCD cameras of the small end-face vision inspection station is shown in Fig. 5. In this figure, the height of the ring light source is adjusted to three levels, namely, (a) high, (b) medium, and (c)

Table 1
Specifications of vision inspection machine components.

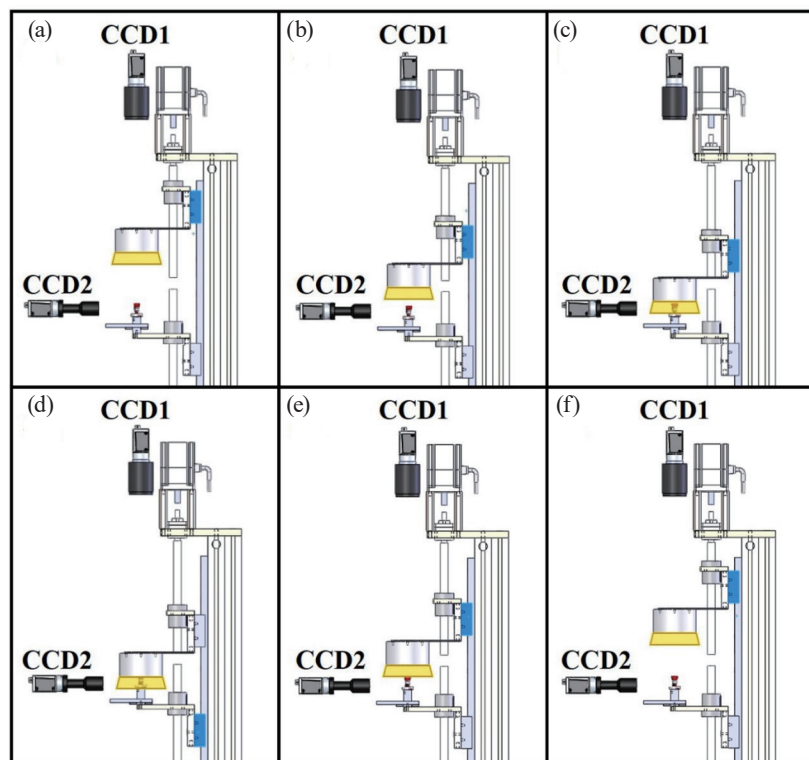| Component	Manufacturer	Specification
Charge-coupled device (CCD)	Basler	acA1300-75gm
Lens 1 of CCD1	VS Technology	50 mm TV lens
Lens 2 of CCD2	VS Technology	1.0×65^2 mm fixed ratio lens
Ring light	High Bright Tech	DLR2-45-070-2-W-24V
Light source controller	High Bright Tech	PC-24V60W-4-S
Operating system	Microsoft	Windows 7 32 bit

Fig. 5.   (Color online) Schematic diagram for height adjustment of light source in small end-face vision inspection station.

low. CCD1 captures images of the top of component 506; after completing the above procedure, CCD2 is used to capture images of the side of the component. Then the height of the ring light source is adjusted to three levels, namely, (d) low, (e) medium, and (f) high, and CCD2 captures images of component 506 at different focal distances. In this study, we only use the images captured by CCD1 for model discussion and thus only three images are processed.

2.1.3   Description of training and recognition equipment

The specifications of the model training equipment used in this study are listed in Table 2, mainly including two NVIDIA TITAN V graphics cards to speed up the training process of the model. The specifications of actual recognition equipment on the production line are shown in Table 3. These specifications are for the model of respective inspection stations performing inspection for the defects of component 506. Therefore, the memory is designed to be relatively large to support the GPU and multithreading to accelerate the calculation in actual execution.

2.2   Image data preprocessing

A total of 430×3 images are initially collected; here, 3 are the images collected at three different focal distances. As the images collected at three different focal distances must be considered, in this research, we adopt the approach of image fusion to reduce the burden on the model. In the fusion approach, three images are fused and then the total number of images is divided by three, so the number of images at this time is 430. Among the 430 images, 274 showed no defects and 156 showed defects. We then perform data augmentation to reduce the validation loss during the training and the overfitting problem, which is particularly useful for small sample data.[21,22] However, in this study, we discuss the model design with only a few samples; thus, we only use random rotation and rescaling. We also consider the balance between defect-positive and defect-negative samples. After the samples are balanced, a total of 548 images are used as data sets for image training, in which there are 274 defect-positive and defect-negative images. Moreover, the recognition rates of DarkNet-53 and AlexNet are compared with that of DenseNet201 in order to retain the meaning of the original model and unify the standard. Therefore, the 548 images are transformed into the corresponding models of DarkNet-53 (25×256), AlexNet (227×227), and DenseNet201 (256×256).

Table 2
Computer specifications of training model.

Type	Model No.
CPU	Intel i9-9820X
Random access memory	8 GB DDR4-2666 x4
Storage Capacity	1T/M.2 PCIe 2280 HDD 4 TB
Graphics card	NVIDIA TITAN V x2

Table 3
Computer specifications of model for actual execution.

Type	Model No.
CPU	INTEL i9-9900K
Random access memory	16 GB DDR4-2666 x4
Storage capacity	512 G/M.2 PCIe 2280 HDD 1 TB
Graphics card	ROG-STRIX-RTX2060-O6G-GAMING

2.3 Network structure improvement

This research mainly utilizes DarkNet-53. The two most important parts of DarkNet-53 are the convolution layer and Resnet. The convolution layer is for extracting features or compressing the number of features to reduce the amount of calculation and the number of parameters for the model, and Resnet is used to solve the degradation problem in the deep network. From experience, we found that the depth of the network has a significant impact on the performance of the model. When the number of network layers is increased, the network can perform and extract more complex feature patterns. Therefore, better results can be achieved with deeper network models in theory, but the accuracy of the network becomes saturated or even decreases with a deeper network, which is the degradation problem. Resnet uses the residual learning method to solve the degradation problem. The accumulation layer only performs identity mapping at this time even when the residual is 0, which is equivalent to a shortcut connection in the circuit; therefore, Resnet can inhibit the decrease in accuracy due to multiple layers of the deep network. In general, the residual value is smaller and the level of learning difficulty is lower; thus, the speed is expected to be higher. In summary, Resnet can make the network deeper, faster, and easier to be optimized, with fewer parameters and lower complexity. The Resnet element added in DarkNet-53 has deepened the network. The network has 53 convolution layers after the improvement and its structure is shown in Fig. 6. DarkNet-53 is mainly used not only for feature extraction in the YOLO V3 model, but also

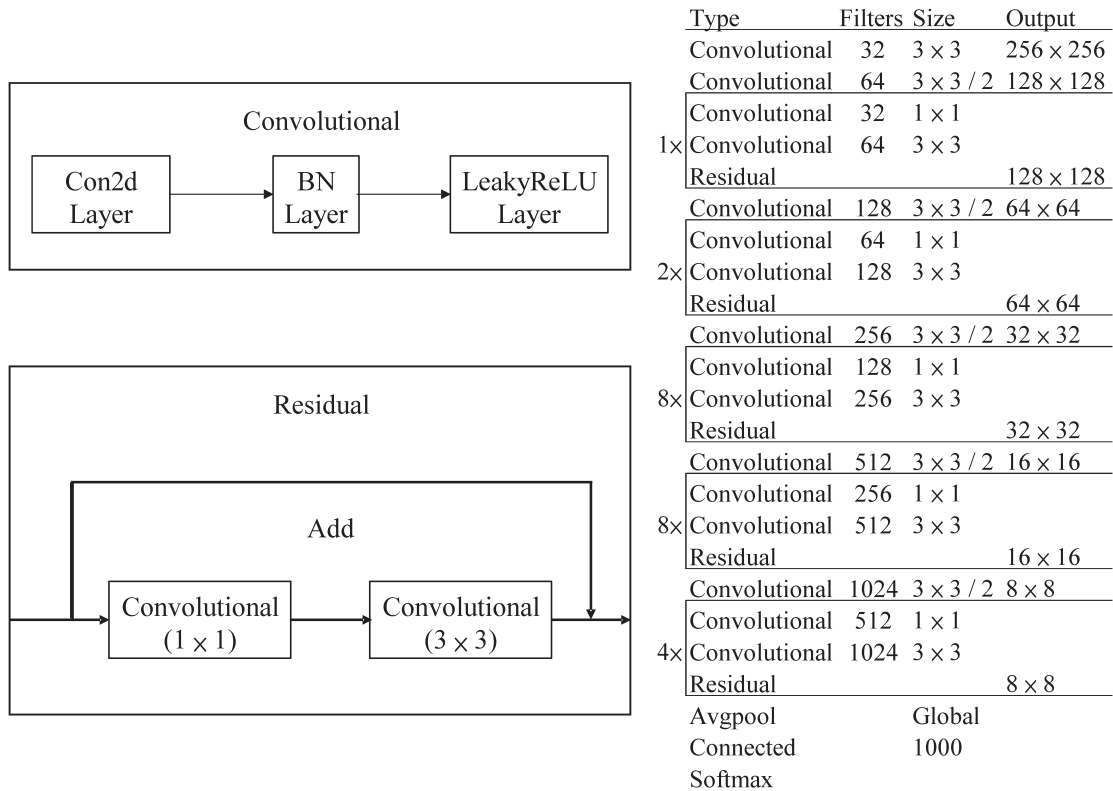|	Type	Filters	Size	Output
	Convolutional	32	3×3	256×256
	Convolutional	64	$3 \times 3 / 2$	128×128
	Convolutional	32	1×1	
1×	Convolutional	64	3×3	
	Residual			128×128
	Convolutional	128	$3 \times 3 / 2$	64×64
	Convolutional	64	1×1	
2×	Convolutional	128	3×3	
	Residual			64×64
	Convolutional	256	$3 \times 3 / 2$	32×32
	Convolutional	128	1×1	
8×	Convolutional	256	3×3	
	Residual			32×32
	Convolutional	512	$3 \times 3 / 2$	16×16
	Convolutional	256	1×1	
8×	Convolutional	512	3×3	
	Residual			16×16
	Convolutional	1024	$3 \times 3 / 2$	8×8
	Convolutional	512	1×1	
4×	Convolutional	1024	3×3	
	Residual			8×8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Fig. 6.   DarkNet-53 structure.

for determining the classification probability of the samples through softmax at the end of the process. In DarkNet-53, the leaky ReLu function is used as the activation function, the pooling layer is discarded, and a convolution layer with a stride of 2 is used to prevent data loss. The network has 53 layers and its nonlinearity increases with depth, processing more spatial features as well as increasing the feature diversity. DarkNet-53 performs a total of five dimensionality reductions, and the numbers of rows and columns of the respective feature matrix of each output for dimensionality reduction decrease to half.

Although Resnet can solve the degradation problem, the accumulation layer only performs identity mapping at this time when the residual is 0. Therefore, the layer upholds its performance, but it still has redundant computations. In addition, the kernel size of the convolution layer can also determine the capability of feature extraction of samples. In summary, in this research, we propose the use of entropy to determine the kernel size of the convolution layer, as well as to determine whether the number of layers has a redundant issue. Shannon referred to and applied the entropy of thermodynamics to information theory, and defined the degree of information variation, which is called information entropy. The entropy indicates how much information is contained in an event, which is called self-information, and it is represented by Eq. (1),[23,24] where p_i represents the probability that event i will occur for the self-information i of event I_i.

$$I_i = -\log_2 p_i \tag{1}$$

The calculation formula for the average information entropy uses the Shannon entropy defined as

$$H = -\sum_{i=0}^{n} p_i \log_2 p_i, \tag{2}$$

which can be derived in grayscale images and rewritten as Eq. (3), where N_i is the number of pixels with image intensity i, and N_s is the total number of pixels for all images.

$$p_i = \frac{N_i}{N_s} \tag{3}$$

In this study, we propose the use of the aforementioned entropy calculation method to determine the kernel size of the convolution layer, as well as to determine whether the number of layers has a redundant issue. The determination process is shown in Fig. 7. The automation process determines the optimal kernel size of the convolution layer. This determination process is based on the difference in entropy among the image samples. The most suitable kernel size of each convolution layer is determined sequentially from the lower layer to the higher layer, and the number of layers can be verified whether it is redundant. The gradient descent with momentum is used in the model training methods in the following experiments.
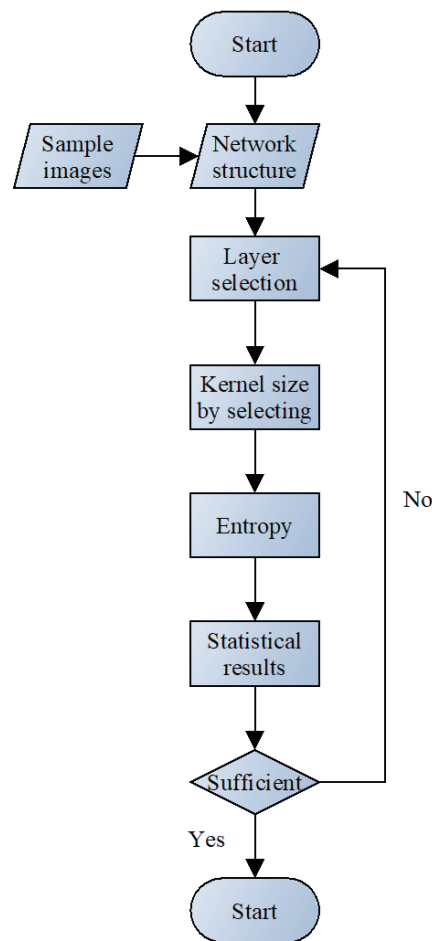
Fig. 7.   Process of determining entropy.

3.   Results and Discussion

3.1   Recognition results of different models with only a few samples

DarkNet-53, AlexNet, and DenseNet-201 are used for training and testing comparison of recognition results of different models with only a few samples. The training parameters of DarkNet-53 are as follows: batch size of 64, 80000 iterations, learning rate of 0.1, momentum of 0.9, weight decay of 0.0005, and an activation function that adopts the leaky ReLU. The training parameters of AlexNet are as follows: batch size of 128, 80000 iterations, learning rate of 0.01, momentum of 0.9, weight decay of 0.0005, and an activation function that adopts ReLU (based on the original model setting). The training parameters of DenseNet201 are as follows: batch size of 128, 160000 iterations, learning rate of 0.1, momentum of 0.9, weight decay of 0.0005, and an activation function that adopts the leaky ReLU. During model training, the data sets are randomly allocated, and 70% of them are used as training sets, 15% as verification sets, and 15% as testing sets. The ratios of samples in each category for the training, verification,

and testing sets are equal. The results of the model comparison are shown in Table 4. For DarkNet-53, the recognition rate is 0.58, the precision is 0.56, the recall is 0.79, and the F1 score is 0.33; for AlexNet, the recognition rate is 0.5, the precision is 0.5, the recall is 1, and the F1 score is 0.33; for DenseNet201, the recognition rate is 0.51, the precision is 0.5, the recall is 0.9, and the F1 score is 0.33. The convergence statuses of the DarkNet-53, AlexNet, and DenseNet-201 models are shown in Figs. 8(a)–8(c), respectively. DarkNet-53 tends to have a more obvious loss of convergence than the other models. The residual learning mechanism of
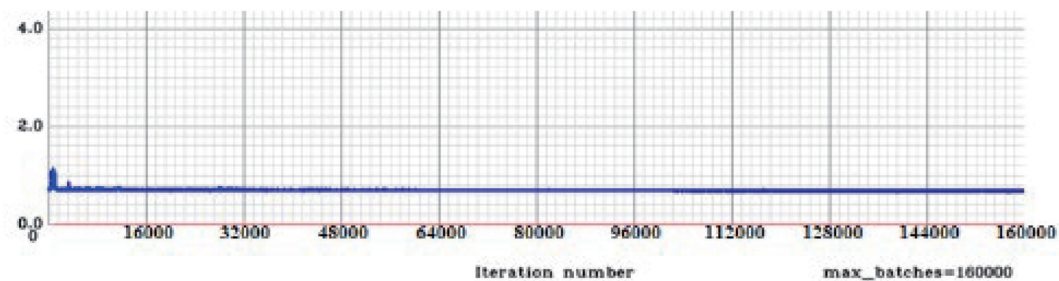
Table 4
Results of various well-known models.

Model	Recognition rate	Precision	Recall	F1 Score
DarkNet-53	0.58	0.56	0.79	0.33
AlexNet	0.50	0.50	1	0.33
DenseNet-201	0.51	0.50	0.99	0.33
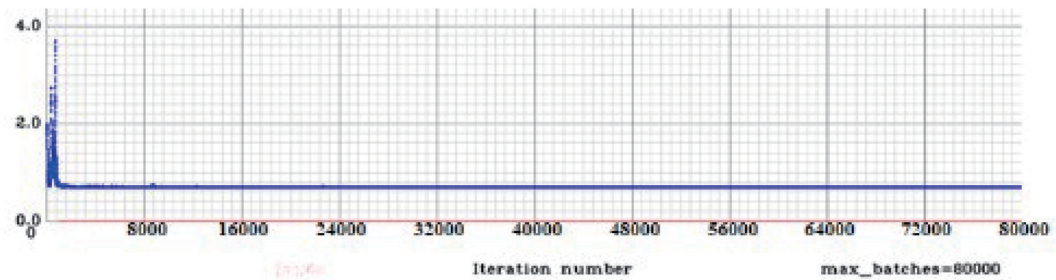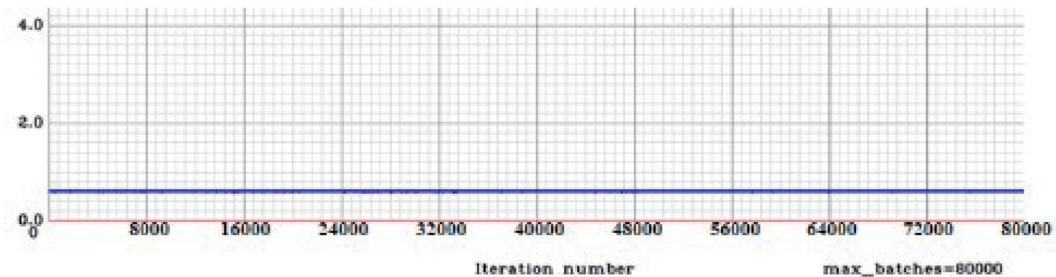
Fig. 8. (Color online) Convergence statuses of three models during training. (a) Convergence status of DarkNet-53, (b) convergence status of AlexNet, and (c) convergence status of DenseNet201.

DarkNet-53 captures and retains the best image feature information, which can more effectively distinguish slight differences between the normal and defective samples of the centralized data. In contrast, without a residual learning mechanism, AlexNet is unable to converge and has the worst performance in terms of recognition rate. DenseNet-201 has a residual learning mechanism similar to that of DarkNet-53; their difference is that DenseNet-201 exhibits the residual learning and serial connection of feature layers in any two convolution layers and any two dense blocks. The closely connected and dependent stacked-layer structure makes DenseNet-201 unable to distinguish the core features between the normal and defective samples as inferred using other unimportant features, affecting the final determination process. Thus, the recognition rate of DenseNet-201 is not higher than that of DarkNet-53.

3.2 Convolution layer using different kernel sizes

The experimental flow processes in Figs. 1 and 7 are carried out to prove the effectiveness of adjusting the size of the convolution layer with the entropy proposed in this study. The best model input image size is obtained to be 192 × 160 by the determination approach using DarkNet-53 with the following training parameters: batch size of 64, 80000 iterations, learning rate of 0.1, momentum of 0.9, weight decay of 0.0005, and an activation function that adopts the leaky ReLU. The experiment for the size of the standardized convolution layer and the experiment after the adjustment of entropy were carried out. The results of these experiments indicate that the recognition rate before the adjustment is 0.58, the precision is 0.56, the recall is 0.79, and the F1 score is 0.33; the recognition rate after the automatic adjustment of entropy reaches 0.81, the precision is 0.74, the recall is 0.95, and the F1 score is 0.42 as shown in Table 5. The obtained convergence statuses before and after adjustment are shown in Figs. 9(a) and (b), respectively. With the entropy adjustment in terms of the size of the convolution layer, Darknet-53 conducts counting for convolution kernel and entropy using statistical centralized data of normal and defective samples, to construct DarkNet-53 with the largest entropy difference in each layer and optimize the model in terms of the effect of feature extraction.

Regarding training time, the training time of DarkNet-53 before the adjustment is 10 h and 32 min. After the adjustment in terms of the size of the convolution layer with entropy, the training time is 4 h and 12 min. The DarkNet53 structure with entropy adjustment in the convolution layers without a difference in entropy adopts the 1 × 1 convolution layer for acceleration in the training, and its number of parameters is smaller than that of the standard DarkNet-53 structure. This mechanism can significantly increase the training speed of models without affecting their accuracy under the condition that the models have the same number of layers.

Table 5
Results of model comparison.

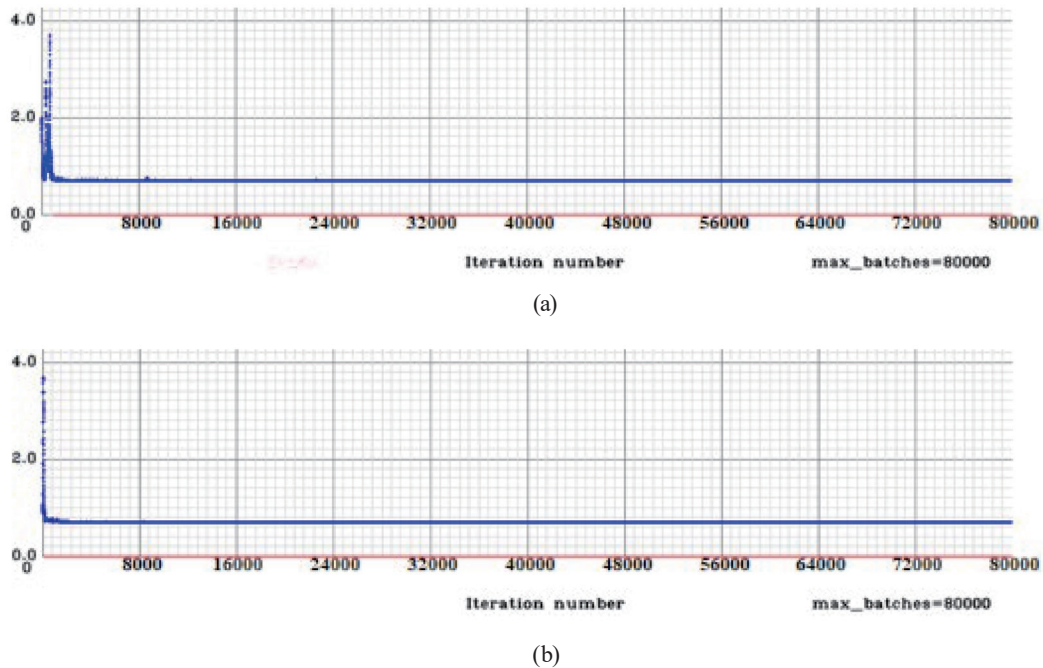| Model	Recognition rate	Precision	Recall
With entropy adjustment DarkNet-53	0.81	0.74	0.95
Standard DarkNet-53	0.58	0.56	0.79

(a)

(b)

Fig. 9.   (Color online) Convergence statuses obtained before and after improvement by method proposed in this study. (a) Convergence status before adjustment and (b) convergence status after adjustment.

4.   Conclusions

In this paper, we mainly discuss how to determine the defects of the component 506 using the DL model with a few samples. We also used an image sensor (ex: camera) to take pictures of the red dotted area of the component. We propose a method of calculating entropy to choose the most suitable kernel size of the convolution layer to enhance the recognition rate of components and shorten the model training time. The experimental results have proved that DarkNet-53 is better than AlexNet and DenseNet-201 with only a few samples. DarkNet-53 has a more obvious convergence tendency in the loss curve and its residual learning mechanism can capture and retain the best image feature information, which effectively distinguishes slight differences in data between normal and defective samples. The images obtained under the influence of the convolution layer with different kernel sizes based on the calculated entropy effectively show an increase in recognition rate and a reduction in training time. The training time in the standard DarkNet-53 model is 10 h and 32 min, and the recognition rate is 58%. After using the method proposed in this study, the training time is 4 h and 12 min, and the recognition rate is 81%. Taken together, the method of selecting the kernel size of the convolution layer proposed in this study can effectively enhance the recognition effect and improve the training time with only a few samples.

Acknowledgments

References

1   G. E. Hinton and R. R. Salakhutdinov: Science **313** (2006) 504. https://doi.org/10.1126/science.1127647
2   Y. LeCun, Y. Bengio, and G. Hinton: Nature **521** (2015) 436. https://doi.org/10.1038/nature14539
3   J. Redmon and A. Farhadi: arXiv preprint arXiv: 1612.08242 (2016).
4   M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, and M. J. Devin: arXiv preprint arXiv:1603.04467 (2016).
5   N. Ketkar: Introduction to Pytorch: Deep Learning with Python: A Hands-on Introduction (Apress, Berkeley, CA, 2017) 1st ed., Chap. 5. https://doi.org/10.1007/978-1-4842-2766-4_12
6   J. Kang, Y. Park, J. Lee, S. Wang, and D. Eom: IEEE Trans. Ind. Electron. **65** (2018) 4279. https://doi.org/10.1109/TIE.2017.2764861
7   W. B. Jiang, Y. Ma, B. Liu, H. K. Liu, B. B. Zhou, J. Zhu, S. Wu, and H. Jin: ACM Trans. Archit. Code Optim. **16** (2019) 23. https://doi.org/10.1145/3357238
8   P. Keserwani, T. Ali, and P. P. Roy: Int. J. Mach. Learn. Cybern. **10** (2019) 3485. https://doi.org/10.1007/s13042-019-00938-1
9   J. Redmon and A. Farhadi: arXiv preprint arXiv 1804.02767 (2018).
10 K. He, X. Zhang, S. Ren, and J. Sun: Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR, 2016) 770. https://doi.org/10.1109/CVPR.2016.90
11 J. Redmon, S. Divvala, R. Girshick, and A. Farhadi: Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR, 2016) 779. https://doi.org/10.1109/CVPR.2016.91
12 K. Simonyan and A. Zisserman: arXiv preprint arXiv 1409.1556 (2014).
13 J. Redmon and A. Farhadi: arXiv preprint arXiv 1804.02767 (2018).
14 D. V. Sang and D. V. Hung: Proc. 10th Int. Symp. Information and Communication Technology (ICPS, 2019) 280. https://doi.org/10.1145/3368926.3369691
15 P. Le, S. Guo, J. Chen, and J. J. Lien: Proc. 2019 Int. Conf. Technologies and Applications of Artificial Intelligence (TAAI, 2019) 1. https://doi.org/10.1109/TAAI48200.2019.8959827
16 A. Ciuparu, A. Nagy-Dabacan, and R. C. Muresan: Neurocomputing **384** (2020) 376. https//doi.org/10.1016/j.neucom.2019.12.014
17 Z. Chen and P. H. Ho: Pattern Recognit. **96** (2019) 10. https//doi.org/10.1016/j.patcog.2019.07.006
18 A. Beke and T. Kumbasar: Eng. Appl. Artif. Intell. **85** (2019) 372. https://doi.org/10.1016/j.engappai.2019.06.016
19 X. Wang, Y. Qin, Y. Wang, S. Xiang, and H. Z. Chen: Neurocomputing **363** (2019) 88. https://doi.org/10.1016/j.neucom.2019.07.017
20 B. O. Ayinde, T. Inanc, and J. M. Zurada: Neural Netw. **118** (2019) 148. https//doi.org/10.1016/j.neunet.2019.04.021
21 S. Mutasa, P. D. Chang, C. Ruzal-Shapiro, and R. Ayyala: J. Digit. Imaging **31** (2018) 513. https//doi.org/10.1007/s10278-018-0053-3
22 H. Ohno: Soft Comput. **24** (2020) 7999. https://doi.org/10.1007/s00500-019-04094-0
23 M. Parker: Digital Signal Processing 101: DCT, Entropy, Predictive Coding, and Quantization (Newnes, Oxford, 2017) 2nd ed., Chap. 24. https://doi.org/10.1016/B978-0-12-811453-7.00024-X
24 M. Welk: Graph Entropies in Texture Segmentation of Images: Mathematical Foundations and Applications of Graph Entropy (Wiley-VCH, Weinheim, 2016) 1st ed., Chap. 7. https://doi.org/10.1002/9783527693245.ch7

About the Authors

Bo-Lin Jian received his BS degree from the Department of Electrical Engineering, National Formosa University in 2009 and his MS degree in materials science and engineering from the National Taiwan University of Science and Technology in 2011. His Ph.D. degree was awarded by the Department of Aeronautics and Astronautics, National Cheng Kung University, in 2017. He is presently an assistant professor in the Department of Electrical Engineering, National Chin-Yi University of Technology, Taichung, Taiwan. His current research interests include signal and image processing, machine learning, and control systems.
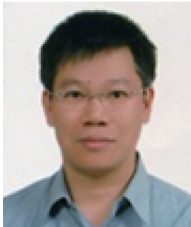
Jui-Pin Hung received his Ph.D. degree from National Chung Hsing University, Taiwan. Currently, he is a professor at the Graduate Institute of Precision Manufacturing at Chin Yi University of Technology in Taichung City, Taiwan. His research interests are in machine tool design, the failure analysis of mechanical components, and orthopedic biomechanics.

Cheng-Chi Wang received his BS, MS, and Ph.D. degrees from the Department of Mechanical Engineering, National Cheng Kung University, Tainan, Taiwan, in 1996, 1998, and 2001, respectively. Now, he serves as a distinguished professor at the Graduate Institute of Precision Manufacturing at National Chin-Yi University of Technology in Taiwan. His current research involves intelligent machining and manufacturing, nonlinear dynamic analysis, simulation and optimization, air lubrication systems, and signal processing. (wcc@ncut.edu.tw)

Chun-Chang Liu received his EMBA degree from Feng Chia University, Taiwan in 2016. Research interests include automobile industry and image processing.