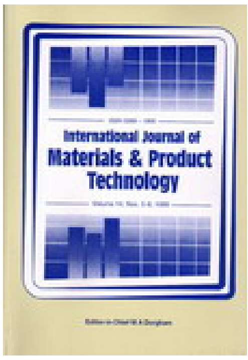


International Journal of Materials and Product Technology

ISSN online: 1741-5209 - ISSN print: 0268-1900
https://www.inderscience.com/ijmpt

Study on detection of dent defects of polariser based on deep convolutional generative adversarial network

Pengfei Shi

DOI: 10.1504/IJMPT.2024.10062361

Article History:

Received:	06 July 2023
Last revised:	04 September 2023
Accepted:	06 November 2023
Published online:	22 February 2024

Study on detection of dent defects of polariser based on deep convolutional generative adversarial network

Pengfei Shi

CETC Fenghua Information-Equipment Co., Ltd.,
Taiyuan, 030024, Shanxi, China
Email: Shipf_1988@163.com

Abstract: The existing techniques of polariser detection only concern whether the polarisers have defects or not and do not classify them as specialised. In addition, lightweight CNN architectures proposed for defect classification of polarisers are based on limited samples. In order to attack the aforementioned issues, a novel grating imaging mechanism based on an adsorption transport platform is designed for a certain defect, dent. Multi-scale negative samples with dent defects and positive samples with other defects or not are expanded by a deep convolutional generative adversarial network (DCGAN). O sets, *64_10000* sets and *128_10000* sets (referred to as the original data, 64*64 generated data and 128*128 generated data) are trained on multiple convolutional neural networks (AlexNet, VGGNet, GoogLeNet, ResNet, SqueezeNet, MoblieNet, ShuffleNet) respectively, the obtained models are then validated on two new samples. Empirically show that ResNet obtained by *64G+128G* perform better than others, classification accuracy rate of the new model is up to 94.94%.

Keywords: deep learning; polariser defect detection; convolutional neural network; CNN; generative adversarial net; GAN.

Reference to this paper should be made as follows: Shi, P. (2024) 'Study on detection of dent defects of polariser based on deep convolutional generative adversarial network', *Int. J. Materials and Product Technology*, Vol. 68, Nos. 1/2, pp.18–28.

Biographical notes: Pengfei Shi received his Master's in Mechanical Engineering from School of Mechanical Engineering of Beijing Institute of Technology in 2013. Currently, he is a Mechanical Engineer in CETC Fenghua Information-Equipment Co., Ltd. His main research interests include deep learning, machine vision inspection technology and system, and image processing.

1 Introduction

A polymeric polariser is one of the most crucial parts of thin-film transistor (TFT) liquid crystal display (LCD) panels, which has a wide range of applications such as car monitor, TVs, wearable devices, mobile phone, pad and computer. Typically, a polariser consists of six transparent layers of polymeric film. During manufacturing, these layers may lead to aesthetic defects such like marks, bubbles, impurities, stains, dents, scratches and so on. The polariser accounts for the total cost of a panel up to 10%, its aesthetic defects play a crucial role on the panel's quality. Almost all kinds of products must be inspected before packing. Thanks to the rapid progress of machine vision and deep learning technology, automated efficient image classification techniques with high precision and speed would be employed to detect these defects above.

But investigation on defects detection of polariser are barely enough recently. Lai et al. (2016) employed a machine vision system which used a LCD monitor to produce a binary stripe pattern for the enhancement of defects imaging. Yen and Syu (2015) developed a cost-effective optical detection system for the tiny bump defects of polariser films based on computer vision techniques. The system applies digital fringes and image processing methods for detecting the tiny bump defects (Yen and Syu, 2015). Deng et al. (2017) proposed a novel automatic inspection method via the structured lighting technique to enhance transparent defects imaging in a polariser. Deng et al. (2018) also proposed an automated inspection method of using edges of light for inspecting ESTADs. They can be detected at the edge of light regions via structured light illumination, which was found to greatly enhance the image contrast (Deng et al., 2018).

Liu et al. (2020a) constructed a lightweight efficient deep learning network for defects classification of polarisers, named LWEN. The shunt module building block and a global average pooling layer after convolutional layer were developed to design the LWEN, which contribute to simplify without reducing the accuracy of classification. Liu et al. (2020b) used a parallel module to develop a real-time detection network for polariser defects based on deep learning, named DNN. Empirically shows that the classification of speed, accuracy and memory utilisation can meet the real-time requirements of industrial production line. Lei et al. (2018) proposed a deep learning technique to train models and detect defects of polariser images based on faster R-CNN, which help to mark the location of defects exactly.

Previous work mainly focused on detecting whether the polarisers have defects or not, and lacked of detailed description of a certain defect detection, that the industry really cares about. Consequently, two significant points were neglected by former researcher. Firstly, the imaging results of different defects under uniform illumination environment was neglected. Because some defects like dent may be observed clearly only in certain imaging condition. Unified lighting conditions cannot perfectly show all kinds of defects, and even may lead to serious erroneous judgment (for some defects that cannot be imaged, the algorithm fails completely). Secondly, the influence of weak supervision of limited samples was ignored. Negative and positive samples (positive samples does not mean no defect) are difficult to collect and algorithm based on small samples are insufficient to meet further defection requirements. Researchers developed some lightweight convolutional neural network (CNN) architectures through studying limited samples, and the generalisation ability under further image samples have not been verified. Unconvincing generalisation ability may also lead to erroneous judgment (OK become NG and NG become OK).

In this paper, we make the following contributions:

- We design a novel grating imaging mechanism based on an adsorption transport platform, which is inspired by the idea of structured light illumination from Lai et al. (2016). The characteristics of dent defects can be clearly captured and other defects may be weakened in some extent.
- We use a deep convolutional generative adversarial network (DCGAN) to expand negative and positive samples and a richer training set are obtained.

- We use a multi-scale generation strategy to expand the samples, and test on multiple CNNs (AlexNet, VGGNet, GoogLeNet, ResNet, SqueezeNet, MoblieNet, ShuffleNet), which empirically show that the training set expanded with 64*64 and 128*128 scale samples perform the best accuracy.

2 Related work

In the past two decades, the increasing development of deep CNNs have make a huge contribution to multiple computer vision tasks such as image classification, object detection, segmentation and localisation. In 2012, Krizhevsky et al firstly proposed a CNN architecture named AlexNet, which won the first place on the ImageNet 2012 classification benchmark (Krizhevsky et al., 2012). Their model have five convolutional layers (some of which are max-polling layers followed) and three fully-connected layers, which lays the foundation for the following architecture. For improving the classification accuracy, researcher try to increase the depth of networks during that period. In 2014, Simonyan proposed VGG network with 16 layers, and Szegedy proposed GoogLeNet with 22 layers (Szegedy et al., 2015; Simonyan and Zisserman, 2014). In VGG networks, they pushed the depth of network to 16–19 through using a smaller convolutional kernel (3×3) architecture than (Krizhevsky et al., 2012), which achieved a significant improvement on classification accuracy. GoogLeNet developed a different module called Inception, the computing resources inside the network is improved utilisation while increasing the depth of the network. He et al. (2016) presented a residual learning architecture to facilitate the training of deeper neural networks, named ResNet. On the ILSVRC 2015, the ResNet with a depth of 152 layers, eight times deeper than VGGNet and about 20 times deeper than AlexNet, achieved the 1st place on classification task. However, with the increasing depth of network, the size of the model becomes bigger and more complicated. Recent studies found that smaller CNNs architecture may offer equivalent accuracy. Iandola et al. (2015) introduced the fire module to build a new smaller CNN architecture named SqueezeNet, which achieved AlexNet-level accuracy with 50× fewer parameters. Howard et al. (2017) proposed an efficient lightweight network for mobile and embedded vision devices named MobileNet. Zhang et al. (2018) presented an extremely computation efficient CNN architecture named ShuffleNet, which is designed for mobile applications with limitation of computing resources.

Supervised learning with CNNs have been successfully applied in industrial defects detection for years. However, due to the insufficient defect images and the high expensive labelled data in practical production lines, it is difficult to get satisfied classification result.

In 2014, Goodfellow et al. firstly proposed generative adversarial nets (GANs) to generate samples. In 2015, Radford et al. proposed a series of CNNs named DCGANs, which help to bridge the gap between the achievement of CNNs for supervised learning and unsupervised learning. Recently, large numbers of researchers take advantage of GAN, DCGAN and their variants to generate image samples, and achieved convincing results in practice. Lian et al. (2020) employed the generative adversarial networks method to expand the limited datasets of the training image samples for defects detection. Qi et al. (2018) developed a multi-view GAN to generate labelled pearl images automatically, and then use the expanded samples to train the multi-stream CNN, results show that the utilisation of images generated can remarkable improve classification performance. Niu et al. (2020) used a large number of defect-free images to generate defect images, the method called surface defect-generation adversarial network (SDGAN), which takes advantage of GANs.

However, a single network structure does not perform very well in the field of polariser detection. The results of combining multiple networks may be even better. There are not only many kinds of defects in polariser but also difficult to image, so there is almost no technical research on using GAN for defect detection. In next section, we use grating-light illumination system to image and purpose multi-scale generation by DCGAN to generate sample.

3 Approaches

3.1 Grating-light illumination

Some defects of polymeric polarisers are hard to characteriser and image via conventional imaging technique, which is called as convex or concave points in the industry, dents may be one of them. Because of their extremely low contrast, these convex or concave points are the primary pain to detecting defects in polymeric polarisers with existing detection technology. To inspect dent defects, Lai et al. (2016) employed structured-light illumination to enhance the imaging defects, which use a LCD monitor to produce a binary stripe pattern. Inspired by the idea of structured light illumination from, we develop a grating-light illumination technique based on an adsorption transport platform.

Figure 1 Grating imaging mechanism based on an adsorption transport platform

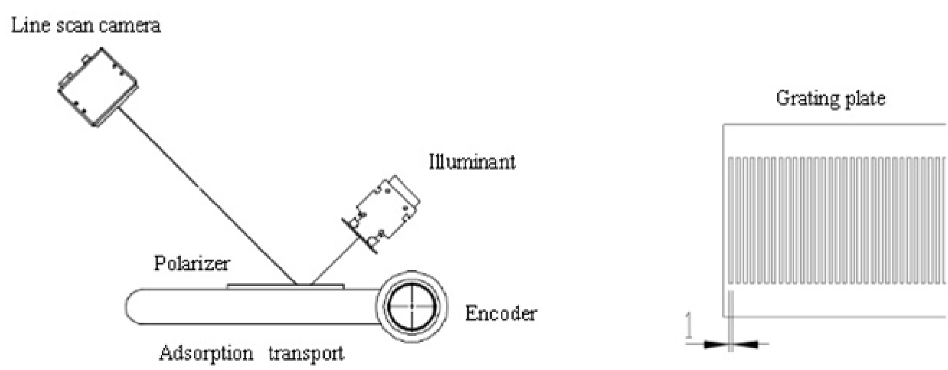


Figure 1 illustrates grating imaging mechanism. Instead of the virtual stripes automatically generated by the visual system, the mechanism uses real stripes generated by grating plate that block light sources, which have higher stability and reliability. The adsorption transport platform can ensure the flatness of polarisers during imaging and meet the demand of on line detection in the industry. In this mechanism, the industrial embedded system utilises the 8k line scan camera take the polariser images based on the position of the polariser captured by the encoder and a sensor trigger. The obtained image data resolution 8,192 × 8,066. Through simple image processing, various 50 × 50 image crops including dent, non-dent and defect-free were extracted. And experienced inspectors classify dent crops into category 1 and the rest into 0. In this paper, the polariser image crops are obtained from three world-famous polariser manufacturing enterprises (short of Se, L, Su) that accounts for more than 90% of the world's production capacity. The material, transparency and refractive index of polariser produced by different manufacturers are various because of different manufacturing process parameters, which may bring about different image quality under the same optical system. Sample details are presented in Table 1.

Table 1 Sample details

Sample	*Category 1*	*Category 0*
Training set (Se)	1,200	2,150
Validation set (Se)	350	525
Test set (L)	565	2,109
Test set (Su)	3,594	4,010

3.2 Multi-scale generation by DCGAN

DCGAN consist of two modelling frameworks (Radford et al., 2015), the generative model G which learn the probability distribution of real data in the training set and transform the input random noise into a image, and the discriminative model *D* which distinguish the fake image generated by model *G* from the real image in the training set. In DCGAN, models *G* and *D* are deep CNNs here, with the parameters represented by θ_g and θ_d, respectively. Define a prior noise z and a training image x, the generator can be represented by a differentiable function $G(z; \theta_g)$ which maps the prior distribution of z to a probability distribution of $G(z)$ in image space. The input of D is x or $G(z)$, and D outputs a single scalar $D(x; \theta_d)$ or $D(G(z); \theta_d)$. Updating G and D simultaneously and the final objective is to obtain the parameters of G. Loss function is with the form below:

$$Min_D \max_G V(D, G) = E_{x \sim p_{data}(x)}\left[\log D(x)\right] + E_{z \sim p_z(z)}\left[\log\left(1 - D\left(G(z)\right)\right)\right] \tag{1}$$

where $p_{data}(x)$ represents the distribution of training image x and $p_z(z)$ represents the distribution of the random noise z.

To make the most use of DCGAN and obtain large scale expanded data, we add a deconvolution layer to generate 128 × 128 scale images based on the model from Radford et al. (2015). More scale means more image diversity, which can show more uncertain details. The generated experimental results are sufficient to support subsequent studies. The overall architecture consists of 6 4 × 4 deconvolution layers and details are depicted in Table 2. ConvT refer to deconvolution operation, which is first discussed for generative model in Radford et al. (2015). Batch normalisation (BN) layers which normalise the input to each unit to have zero mean and unit variance are added in ConvT with the exception of ConvT6. The BN layers can deal with poor initialisation problems in training process and helps gradient go in deeper. Every ConvT layers are followed by ReLU nonlinearity layer but CovT6, which is added by Tanh function. In this paper, we use the generation model with the resolution of 64 × 64 and 128 × 128 to generate new training sets with a number of 10,000 (0 and 1 samples account for half respectively) based on the training set in Table 1 (the samples come from different companies, and the defects have been confirmed by several experienced engineers for many times. Table 1 shows the quantity after selection).

Table 2 G model architecture for 128 × 128 scale

Layer	*Kernel size*	*Input channels*	*Output channels*	*Output size*
ConvT1	4×4	100	1024	4×4
BN		1024	1024	4×4
ReLU				4×4
ConvT2	4×4	1024	512	8×8
BN		512	512	8×8
ReLU				8×8
ConvT3	4×4	512	256	16×16
BN		256	256	16×16
ReLU				16×16
ConvT4	4×4	256	128	32×32
BN		128	128	32×32
ReLU				32×32
ConvT5	4×4	128	64	64×64
BN		64	64	64×64
ReLU				64×64
ConvT6	4×4	64	3	128×128
Tanh				128×128

3.3 Experiments details

Firstly, we train eight representative and classical CNN models (AlexNet, VGG16, GoogLeNet, ResNet, SqueezeNet, MoblieNet, ShuffleNet, DenseNet) using stochastic gradient descent with the same batch-size of 32, same momentum of 0.9, and same weight decay of 0.0005 on training sets (Se) for 100 cycles. Initialisation parameters and regularisation strategy follow the original work. The learning rate is initialised at 0.01 and is divided by 10 every 30 cycles. Then the trained models (named model sets I) are tested on the test set (L) and test set (Su).

Secondly, we extend samples based on the training set (Se) and new samples are named as *64_10000*, *128_10000*, where the number in front of the underline refers to the resolution, and the number after the underline refers to the number of samples. The batch-size of the G, D models are all set to 64, and use the Adam optimisation algorithm with the learning rate of 0.0002 for 64 × 64 image generation. And for 128 × 128 image generation, the learning rate is set to 0.00001. The two models are alternately trained until convergence, approximately 700 cycles.

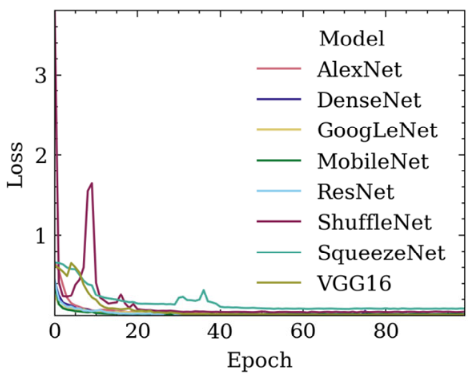
Thirdly, these model sets are trained for transfer learning on various extended sample sets, then the trained models are tested on the test set (L) and test set (Su).

4 Results and discussion

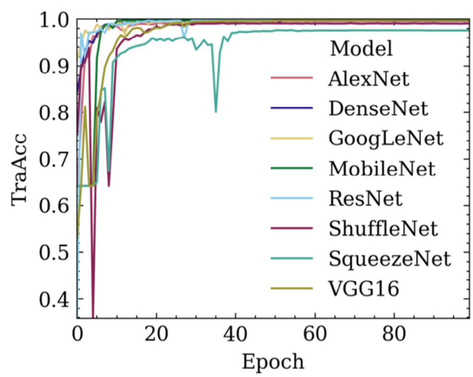
4.1 Results of model sets I

Figure 2 shows the training loss curves, training accuracy curves and validation accuracy curves of the eight models on datasets from Se in Table 1. Figure 2(a) reveals that all the eight classical models can converge quickly and loss values are all converge to 0 after 100 training cycles. Figure 2(b) and Figure 2(c) reveal that the eight classical models perform very well in both the training set (Se) and the validation set (Se). Here, we can conclude that these eight classical models perform perfectly for datasets from the same source. But for different sources, they do not do well.

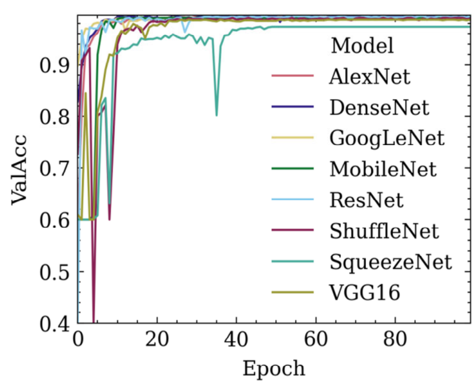
Figure 2 (a) Comparison of training loss curves on the training set (Se) (b) Comparison of classification accuracy curves on the training set (Se) (c) Comparison of classification accuracy curves on the validation set (Se) (see online version for colours)



(a)



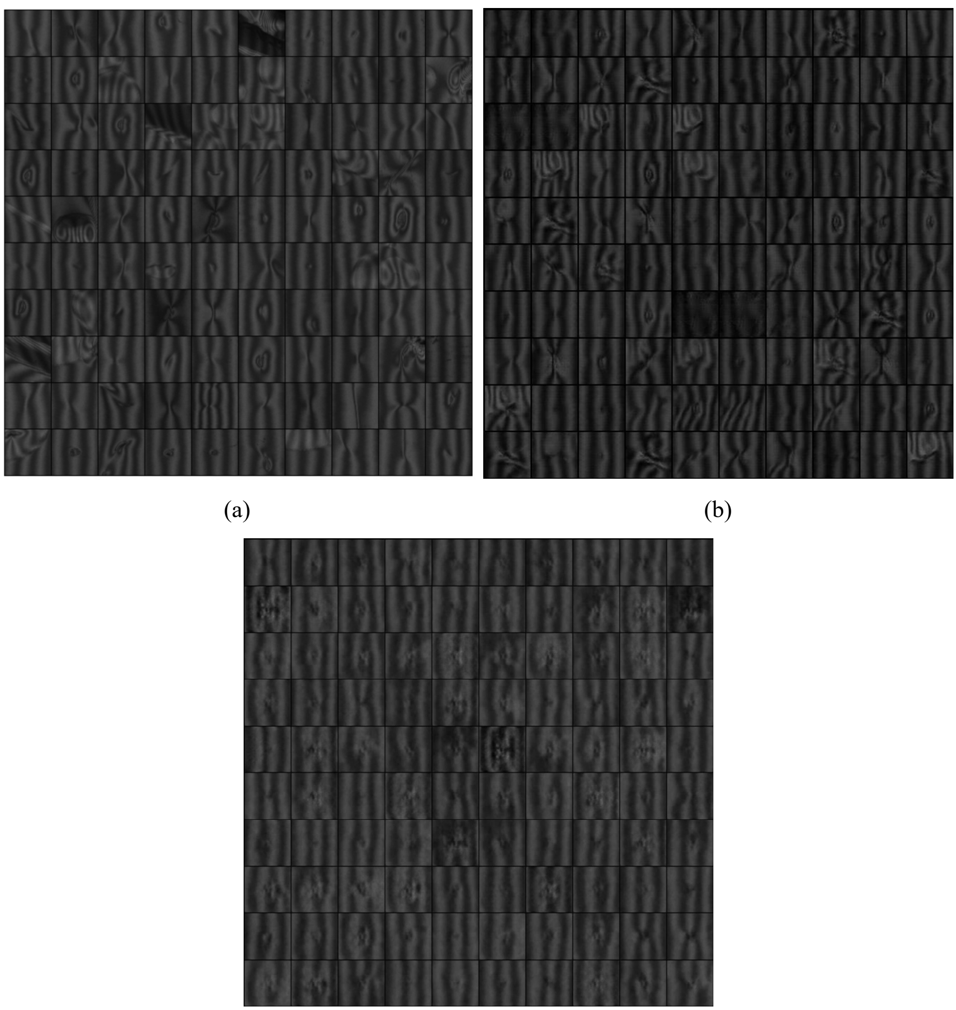
(b)



(c)

The trained eight models are tested on the test set (L) and test set (Su), results are showed in Table 3. Unfortunately, only three models have an accuracy of more than 80% on the test set (L) and test set (Su), they are GoogLeNet, ResNet and DenseNet. For GoogLeNet, the accuracy on test set (L) and Test set (Su) are 81.76%, 85.19% respectively. ResNet are 86.76%, 80.84 respectively. And for DenseNet are 80.88%, 82.07% respectively. Some of the early CNN models such as AlexNet and VGG16 do not perform well may due to their flatten architecture. And lightweight CNNs like SqueezeNet, MoblieNet and ShuffleNet also perform poorly. We argue the lower accuracy is caused by the difference between training sets and test sets. Because of different manufacturing technique and process parameters of the three polariser manufacturing enterprises, the material, transparency and refractive index of polariser produced are various. Therefore, the optical properties of the polariser surface are also different, which play a key role on defects detection. Hence, this study has more practical significance, it contribute to utilise existing datasets to fit unknown datasets that may have different features.

Figure 3 (a) Real images for 50 × 50 from training set (Se) (b) Fake images for 64 × 64 (c) Fake images for 128 × 128



(a) (b)

(c)

The similarity of two images is obtained by calculating cosine distance between image vectors. Tables 4 are the computing results between training set (Se) and other datasets.

Table 3 Comparison of classification accuracy (%) of eight models

Model	*Training set (Se)*	*Validation set (Se)*	*Test set (L)*	*Test set (Su)*
AlexNet	99.61	99.31	49.78	66.89
VGG16	99.85	98.47	45.74	62.87
GoogLeNet	99.91	98.74	*81.76*	*85.19*
ResNet	100	99.43	*80.84*	*86.76*
SqueezeNet	97.61	97.37	54.97	56.68
MoblieNet	99.79	99.09	72.94	76.59
ShuffleNet	99.07	98.97	46.67	60.06
DenseNet	100	99.43	*80.88*	*92.07*

Table 4 Average similarity between datasets (%)

Datasets	*Training set (Se)*	*Validation set (Se)*	*Test set (L)*	*Test set (Su)*
Training set (Se)	96.01	95.81	86.82	91.57

The similarity between training set and validation set are 95.81% because they are from the same dataset. And the similarity between training set (Se) and the test set (L), test set (Su) are 86.82% and 91.57% respectively. Furthermore, image data from test set (Su) are more close to training set and table 3 shows this. Classification accuracy of 8 models on test set (Su) are slightly higher than on test set (L).

We generate datasets based on the training set (Se) named as 64_10000 and 128_10000. Figure 3 illustrates the real images and fake images. The three models that performed well (GoogLeNet, ResNet and DenseNet) are trained on new datasets of 64_10000, 128_10000, 64_10000 + 128_10000 respectively, then they are tested on test set (L) and test set (Su).

4.2 Results of models with expanded datasets

The classification accuracy on the test set (L) and test set (Su) of three models which are trained on original and expanded datasets are showed in Table 5, Table 6, Table 7 respectively. The classification accuracy of the models trained with the expanded dataset on test sets are all improved in some extent (6.68%–13.47%). And the ResNet do the best, the classification accuracy of ResNet trained with training set (Se) +64_10000+128_10000 on test set (L) and test set (Su) are 95.65%,94.25% respectively, which improved about 13.47%.

Table 5 Comparison of classification accuracy (%) of ResNet

Training set	*Test set (L)*	*Test set (Su)*	*Average relative IMP (%)*
Training set (Se)	80.84	86.76	--
Training set (Se)+64_10000	91.36	93.58	10.44
Training set (Se) +128_10000	90.56	91.53	6.68
Training set (Se) +64_10000+128_10000	95.65	94.23	13.47

Table 6 Comparison of classification accuracy (%) of GoogLeNet

Training set	*Test set (L)*	*Test set (Su)*	*Average relative IMP (%)*
Training set (Se)	81.76	85.19	--
Training set (Se)+64_10000	90.87	92.15	9.66
Training set (Se) +128_10000	90.24	92.50	9.48
Training set (Se) +64_10000+128_10000	90.17	92.78	9.60

Table 7 Comparison of classification accuracy (%) of DenseNet

Training set	*Test set (L)*	*Test set (Su)*	*Average relative IMP (%)*
Training set (Se)	80.88	92.07	--
Training set (Se)+64_10000	90.99	93.73	7.15
Training set (Se) +128_10000	90.10	92.88	6.15
Training set (Se) +64_10000+128_10000	91.41	94.24	7.69

5 Conclusions

- The novel grating imaging mechanism based on an adsorption transport platform can effectively captured the characteristics of dent defects and weaken other defects in some extent.
- We use a multi-scale generation strategy to expand the samples, and test on multiple CNNs, which empirically show that the training set expanded with 64*64 and 128*128 scale samples perform the best accuracy.
- ResNet has the best performance among many CNN models, so we mainly use ResNet for defects detection in production practice, and developed defect detection equipment for polariser enterprises.
- With the development of the research work, the number and type of samples will continue to expand, and the single type of defect generation strategy has gradually failed to meet the production needs, so the generation strategy of multi-classification output is the research focus of the next step.

References

Deng, Y-l., Xu, S-p. and Lai, W-w. (2017) 'A novel imaging-enhancement-based inspection method for transparent aesthetic defects in a polymeric polarizer', *Polymer Testing*, Vol. 61, No. 13, pp.333–340.

Deng, Y-l., Xu, S-p., Chen, H-q., Liang, Z-h. and Yu, C-l. (2018) 'Inspection of extremely slight aesthetic defects in a polymeric polarizer using the edge of light between black and white stripes', *Polymer Testing*, Vol. 65, No. 26, pp.169–175.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M. et al. (2014) 'Generative adversarial nets', *International Conference on Neural Information Processing Systems*.

He, K-m., Zhang, X-y., Ren, S-q. and Sun, J. (2016) 'Deep residual learning for image recognition', Paper presented at *2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV.

Howard, A.G., Zhu, M-l., Chen, B. et al. (2017) *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*, arXiv preprint arXiv:1704.04861.

Iandola, F.N., Han, S., Moskewicz, M.W. et al. (2016) *Squeezenet: Alexnet-Level Accuracy with 50x Fewer Parameters and < 0.5 MB Model Size*, arXiv preprint arXiv:1602.07360.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) 'ImageNet classification with deep convolutional neural networks', Paper presented at *Neural Information Processing Systems 2012*, Stateline, NV.

Lai, W-w., Zeng, X-x., He, J. and Deng, Y-l. (2016) 'Aesthetic defect characterization of a polymeric polarizer via structured light illumination', *Polymer Testing*, Vol. 53, No. 31, pp.51–57.

Lei, H-w., Wang, B. et al. (2018) 'Defect detection for polymeric polarizer based on faster R-CNN', *Journal of Information Hiding and Multimedia Signal Processing*, Vol. 9, No. 6, pp.1414–1420.

Lian, J., Jia, W., Zareapoor, M. et al. (2020) 'Deep-learning-based small surface defect detection via an exaggerated local variation-based generative adversarial network', *IEEE Transactions on Industrial Informatics*, Vol. 16, No. 2, pp.1343–1351.

Liu, R-z., Sun, Z-y., Wang, A-h. et al. (2020a) 'Lightweight efficient network for defect classification of polarizers', *Concurrency and Computation: Practice & Experience*, Vol. 32, No. 11, p.e5663.

Liu, R-z., Sun, Z-y., Wang, A-h. et al. (2020b) 'Real-time defect defection network for polarizer based on deep learning', *Journal of Intelligent Manufacturing*, Vol. 31, No. 37, pp.1813–1823.

Niu, S., Li, B., Wang, X-g. et al. (2020) 'Defect image sample generation with GAN for improving defect recognition', *IEEE Transactions on Automation Science and Engineering*, Vol. 17, No. 3, pp.1611–1622.

Qi, X., Chen, Z-z., Liu, Y. et al. (2018) 'Multiview generative adversarial network and its application in pearl classification', *IEEE Transactions on Industrial Electronics*, Vol. 66, No. 10, pp.8244–8252.

Radford, A., Metz, L. and Chintala, S. (2015) *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks*, arXiv preprint arXiv:1511.06434.

Simonyan, K. and Zisserman, A. (2014) *Very Deep Convolutional Networks for Large-Scale Image Recognition*, arXiv preprint arXiv:1409.1556.

Szegedy, C., Wei, L., Jia, Y-q. et al. (2015) 'Going deeper with convolutions', Paper presented at *2015 IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA.

Yen, H-N. and Syu, M-J. (2015) 'Inspection of polarizer tiny bump defects using computer vision', Paper presented at *2015 IEEE International Conference on Consumer Electronics*, Las Vegas, NV.

Zhang, X-.u, Zhou, X-y., Lin, M-x. et al. (2018) 'ShuffleNet: an extremely efficient convolutional neural network for Mobile devices', Paper presented at *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT.