Saw-Mark Defect Detection in Heterogeneous Solar Wafer Images using GAN-based Training Samples Generation and CNN Classification

Du-Ming Tsai[1], Morris S. K. Fan[2], Yi-Quan Huang[1] and Wei-Yao Chiu[1]
[1]*Department of Industrial Engineering and Management, Yuan-Ze University, 135 Yuan-Tung Road, Chung-Li, Taiwan, Republic of China*
[2]*Department of Industrial Engineering and Management, National Taipei University of Technology, 1 Sec. 3 Zhongxiao E. Rd., Taipei, Taiwan, Republic of China*

Keywords: Defect Detection, Multicrystalline Solar Wafer, Saw Mark, Deep Learning.

Abstract: This paper presents a machine vision-based scheme to automatically detect saw-mark defects in solar wafer surfaces. A saw-mark defect is a severe flaw when cutting a silicon ingot into wafers. A multicrystalline solar wafer surface presents random shapes, sizes and orientations of crystal grains in the surface and, thus, results in a heterogeneous texture. It makes the automatic visual inspection task extremely difficult. The deep learning technique is an ideal choice to tackle the problem, but it requires a huge amount of positive (defect-free) and negative (defective) samples for the training. The negative samples are generally not sufficient enough in a manufacturing process. We thus apply a GAN-based model to generate the defective samples for training, and then use the true defect-free samples and the synthesized defective samples to train a CNN model. It solves the imbalanced data arising in manufacturing inspection. The preliminary experiment has shown promising results of the proposed method for detecting various saw-mark defects including black line, white line, and impurity in multicrystalline solar wafers.

SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

Solar power has become an attractive alternative of electricity energy in recent years. For the currently available solar cell technologies, multicrystalline solar cells dominate the market share owing to lower manufacturing costs. A main category of defects found in silicon solar wafers is called "saw-mark". It occurs when a silicon ingot is sliced into wafers in the cutting process with a multi-wire sawing technique. This paper presents a machine vision-based scheme to automatically detect saw-mark defects in multicrystalline solar wafers.

A saw-mark defect is a severe flaw of wafers for making solar cells. It contains potential cutting stress that may cause cracks in a thin silicon wafer. It also reduces the power transmission efficiency. Therefore, detection of saw-mark defects in sliced solar wafers at the early processing stage is demanding in solar wafer manufacturing. A multicrystalline solar wafer presents random shapes, sizes and directions of crystal grains in the surface and results in a heterogeneous texture. The textured surface shows local random patterns in the background and, thus, makes the saw-mark defect hardly distinguishable from the faultless regions. Fig. 1(a) shows the image of a defect-free multicrystalline solar wafer surface. It contains multiple grains of random shapes and sizes. Fig. 1(b)-(d) presents three different saw-mark types. Fig. 1(b) is a thick groove that results in a black line saw-mark in the image. Fig. 1(c) is a thin groove and is shown as a white stripe saw-mark in the image. Fig. 1(d) is a saw-mark defect caused by the saw slicing through an impurity.

The surface defects of a solar wafer or a solar cell result in high recovery cost in the manufacturing process and reduction in production yield. This calls for automatic visual inspection of solar wafers/cells. (Fu et al., 2004) implemented a machine vision scheme to detect edge crack of solar cells. It only inspected the cracks in the solar cell edges with obvious gray-level variances. (Ordaz and Lush, 2000) analyzed the converting efficiency of a solar cell based on the gray-level distribution in the electroluminescence image. (Pilla et al., 2002)

Tsai, D., Fan, M., Huang, Y. and Chiu, W.
Saw-Mark Defect Detection in Heterogeneous Solar Wafer Images using GAN-based Training Samples Generation and CNN Classification.
DOI: 10.5220/0007306602340240
In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)*, pages 234-240
ISBN: 978-989-758-354-4
Copyright © 2019 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

applied thermographic inspection of photovoltaic solar cells to identify cracks. Most of the solar cell inspection methods focus on efficiency assessment and edge crack detection, and the surface defects are rarely mentioned. (Tsai et al., 2010) proposed an anisotropic diffusion scheme for detecting micro-crack defects in multicrystalline solar wafers. The micro-crack in the sensed image presents low gray-level and high gradient characteristics. The anisotropic diffusion scheme works successfully for detecting micro-cracks in multicrystalline solar wafers. However, it can not be extended to the detection of saw-mark defects in solar wafer images.

For heterogeneously textured surfaces, similar patterns will not repeatedly appear in the image. To detect defects in a heterogeneous texture such as marbles or granites, (Ar and Akgul, 2008) employed eight Gabor filters to construct a feature extraction system for marble tile inspection. (Xie and Mirmehdi, 2005 and 2007) presented an automatic defect detection method for random color-texture surfaces. It generated a set of texture exemplars by exploring a Gaussian mixture model from defect-free image patches, and used them for defect detection in marble tiles. (Li and Tsai, 2011) proposed a global spectral-domain solution to detect saw-mark defects in multicrystalline solar wafers. The Fourier image reconstruction is used to smooth out the main background pattern. Then, the Hough transform is applied in the reconstructed image to find the saw-marks that are deviated from the Hough-lines. Since the method requires both Fourier transform and Hough transform, it is computationally expensive in the inspection process. It is specially designed for saw-mark defects, and cannot be extended to detect other defect types such as particles and fingerprints.

Deep learning (LeCun et al., 2015) has been a popular and dominant technique in computer vision for object detection and object recognition. It is well suited for industry inspection applications because the end-to-end model requires no handcrafted features. (Soukup and Huber-Mork, 2014) used the CNNs to detect defects in the photometric stereo images of non-textured metal surfaces. (Li et al., 2017) proposed Fisher criterion-based autoencoders to detect local defects in textile fabric. It is applied to the surfaces with homogeneous textures or repetitive patterns. (Cha et al., 2017) also used the CNN to detect crack damage in concrete surfaces. It trained up to 40K images that contain non-textured surfaces. (Gibert et al., 2017) used the CNN to inspect the railway steels. A high recognition rate is reported. The methods above mainly focus on non-textured or homogeneously textured surfaces.

In the manufacturing environment, especially in the product pilot-run stage, it is easy to collect defect-free samples as many as required. However, it is difficult to collect a sufficient number of defective samples in a short period of time. The success of a well-trained deep learning neural network generally depends on a huge number of training samples, where both positive and negative datasets should be roughly the same in size. The proposed deep learning scheme for saw-mark detection in heterogeneous solar wafer images is thus composed of two phases: defect samples generation using the CycleGAN (Cycle-consistent adversarial networks, Zhu et al., 2017), and then defect detection using the CNN (convolutional neural networks, Krizhevsky et al., 2012) based on the true defect-free samples and the synthesized defective samples. This approach allows the CNN to train as many required positive and negative samples as possible to obtain the best inspection result.

The paper is organized as follows. In section 2, the CycleGAN used for defect samples generation is first described. The CNN model used for saw-mark detection is then presented. In section 3, the experimental results on full-sized solar wafer images are analyzed. Section 4 concludes this paper.

2 DEEP LEARNING MODELS

This section presents the machine vision scheme for saw-mark detection in solar wafer images, which includes the GAN-based model for defect samples generation and the CNN model for defect detection. As discussed in the last section and shown in Figure 1, the multicrystalline solar wafer image contains random, irregular crystal grains. The training and inspection cannot use the whole sensed wafer image as the input to detect small local defects. Instead, small image patches are randomly selected from the solar wafer images. The image patches are the input to the CycleGAN and CNN models. In the inspection process, a window of the patch size is slid pixel by pixel over the full inspection image, and is fed individually to the trained CNN model for the classification.

2.1 CycleGAN for Defect Samples Generation

In this study, we use the CycleGAN developed by (Zhu et al., 2017), instead of the GAN (Goodfellow et al., 2014), to generate representative defect samples from a very limited number of true saw-mark defects. The objective of the CycleGAN model combines both

the adversarial loss (just like GAN, Goodfellow et al., 2014) and the cycle consistency loss (Zhou et al., 2016) to create the output images. It measures the adversarial lose for matching the distribution of generated images to the data distribution in the target domain. The consistency lose is used to prevent the learned forward and backward mappings from contradicting each other. It does not use specific paired samples as training data. Instead, it uses unpaired datasets for the training, and is suited for our application. It can capture special characteristics (especially the color and texture) from one image collection and learn how these characteristics can be translated into other image collection without any paired training examples. To generate artifacts of saw-mark defects, we use the true defect patches as the target dataset in the CycleGAN, and then randomly collect a small set of defect-free patches from the solar wafer images as the input set to the CycleGAN model. The CycleGAN will learn the transformation from the input defect-free patches into a set of defective patches. Whenever we change the input set with different defect-free patches to the trained CycleGAN, a new defective set is created. Since we can have as many true defect-free samples as we want, the CycleGAN can create as many synthesized defective samples as we need for the CNN model. The architecture of the CycleGAN model for defect samples generation is illustrated in Figure 2.

Figure 3(a) shows 10 demonstrative true defect-free solar image patches used as the input dataset to the CycleGAN. Figure 3(b) and (c) presents respectively ten true black saw-mark and ten true white saw-mark defect samples used as the target set of the CycleGAN. The real datasets, such as those in Figure 3(a)-(c), are used to train the CycleGAN. The size of the image patch is 50×50. Figure 4(a) displays a set of defect-free samples to the trained CycleGAN, and Figure 4(b) and (c) shows the resulting black and white saw-mark defect samples generated by the CycleGAN. It shows that the synthesized defect patches present similar textured characteristics as those of the true defect sample patches. The defect part in the image does not show clear edge changes from its surroundings, whereas the crystal grain edges are sharp and clear.

2.2 CNN Classification Model and Defect Detection

The CNN model is used for classifying an unknown image patch as defect-free or defective. In this study, a simple CNN that comprises 3 convolutional layers are used for the training. A CNN model with a limited number of convolutional layers gives better computational efficiency in the inspection process. Figure 5 depicts the detailed structure and shows the main parameters of the proposed CNN models. The real defect-free image patches collected from the solar wafers are used as the positive samples and the synthesized defective image patches produced by the CycleGAN are used as the negative samples to the CNN for training. Illumination normalization is applied to both positive and negative samples prior to the CNN training.

In the inspection process, a window of the size of the image patch used in the neural networks is moved pixel by pixel throughout the full-sized multicrystalline solar wafer image. The windowed image patch is then fed to the trained CNN model for classification. The central coordinates of the window will be marked in black in the full-sized image if the image patch is classified as a defect. Reversely, it is marked in white if the patch is classified as a normal one. The resulting black region in the binary image gives the shape and location of a detected defect in the solar wafer surface. Let W(x, y) be the window patch with the center at (x, y) in the full-sized image to be inspected. The resulting binary image is given by

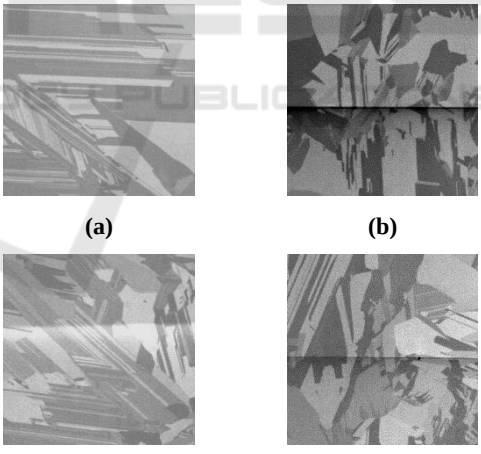

Figure 1: Solar wafer surfaces: (a) defect-free solar wafer image; (b) solar wafer image with a black saw-mark defect; (c) white saw-mark defect; (d) saw-mark defect caused by impurity.

$$B(x,y)=\begin{cases}1\ (black), if\ W(x,y)\in Defect\ detected\ by\ CNN\\ 0\ \ (white), otherwise\end{cases} \quad (1)$$

Since the saw-mark in a small windowed patch contains only subtle changes with respect to the random grain textures, the entire saw-mark region

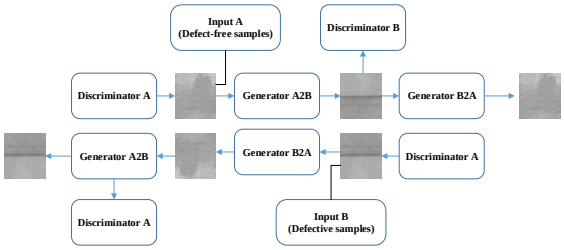

Figure 2: The CycleGAN model used for defect patches generation.

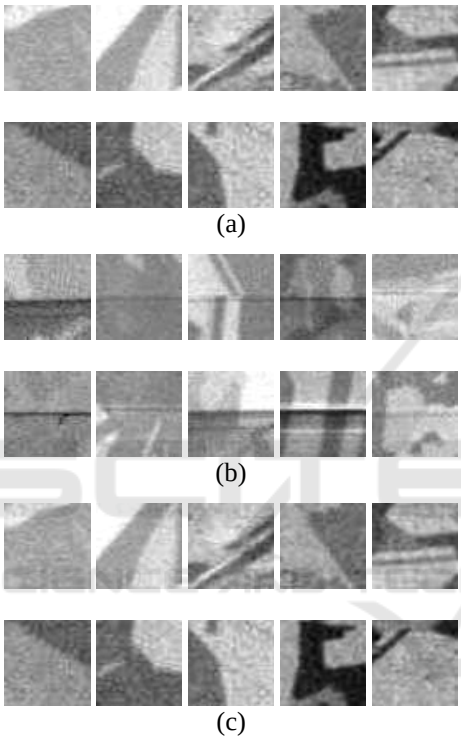

Figure 3: Real solar wafer image patches used for training the CycleGAN model: (a) real defect-free samples; (b) real black saw-mark samples; (c) real white saw-mark samples.

may not be completely detected in the full-sized solar wafer image. We thus further apply the horizontal projection line by line in the resulting binary image B(x, y) to intensify the horizontal saw-mark in the image. That is

$$P(y) = \sum_{x} B(x, y), \forall y \qquad (2)$$

The maximum projection value is then used as the discriminant measure for saw-mark detection, i.e. $P(y^*) = \max\{P(y), \forall y\}$. If the horizontal projection $P(y^*)$ is large enough, a saw-mark at line y^* is declared.

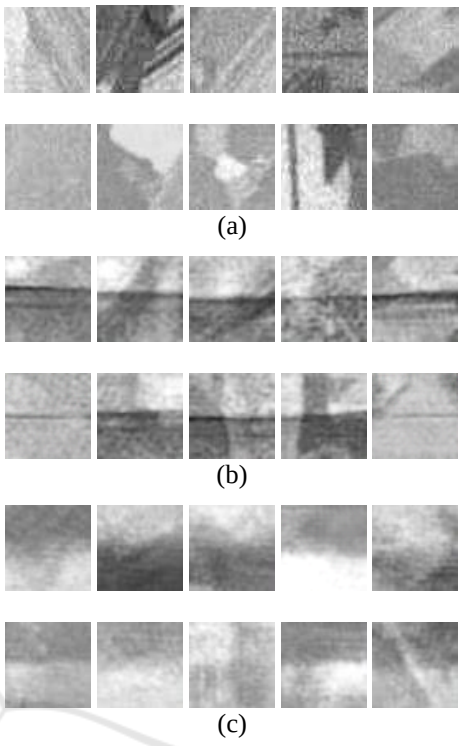

Figure 4: Synthesized defect patches generated by CycleGAN: (a) real defect-free samples input to the trained CycleGAN; (b) generated black saw-mark patches; (c) generated white saw-mark patches.

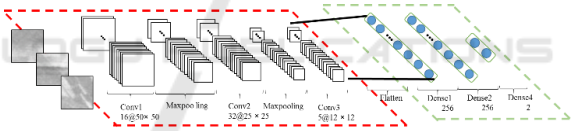

Figure 5: The CNN model used for defect detection.

3 EXPERIMENTAL RESULTS

This section presents the experimental results on a number of solar wafer images containing various saw mark defects to evaluate the performance of the proposed defect detection scheme. The test images are 500×500 pixels wide with 8-bit gray-levels. The window patch is of size 50×50 pixels. All the test images conducted in the experiment are captured from real solar wafer surfaces.

The proposed algorithms were implemented on a personal computer with an Intel Core 2, 3.6GHz CPU and an NVIDIA GTX 1070 GPU. The mean computation time of the proposed method is 0.004 seconds for an image patch of size 50×50 pixels. To train the CycleGAN model for defect samples generation, a small number of 150 real defect-free

patches and 150 real defective patches of the solar wafers are used as the training samples. To train the CNN model for defect classification, a total of 4000 real defect-free patches and 4000 synthesized saw-mark patches are used as the training samples. While training the CNN model, a set of 150 real defect-free patches and a set of 150 real defective patches are used to verify the effectiveness of the trained CNN. The test results show that the FN (missing detection) rate is 22%, and the FP (false alarm) rate is 3%. The final CNN trained is then used to inspect the full-sized solar wafer image by sliding the window pixel by pixel. Since the saw-mark is shown as a long stripe or line across the solar wafer, the recognition rate of 78% in terms of image patches can still reliably detect the presence of a saw-mark defect in the inspection image.

Figure 6(a1)-(a5) shows five defect-free solar wafer images, and (b1)-(b5) illustrates the detection results by superimposing the suspected defect pixels in the original images. The profiles shown in Figure 6(c1)-(c5) are the corresponding horizontal projection P(y). The proposed defect-detection scheme can reliably ignore the normal grain patterns in the detection process and results in clear surfaces in the final binary images.

Figure 7(a1)-(a5) further presents five defective solar wafer images that contain dark and bright saw marks. Some saw marks are very thin and low-contrasted. In Figure 7(a1), there is a horizontal dark stripe without clear edges in the image, and the saw-mark is not distinctly visible. As observed from the projection profiles, the defect-free solar wafer images present very low P(y*) values close to zero, whereas all defective solar wafer images yield distinctly large projection values. A preliminary test on 15 defect-free and 15 defective solar wafer images shows that the proposed method can correctly identify all types of saw-marks without false alarms with a proper threshold setting for P(y*).

The proposed method for defect detection with imbalanced data is also compared with the under-sampling, over-sampling (Chawla et al., 2002, Yen and Lee, 2009) and class weights. Let n^+ be the number of positive (defect-free) samples and n^- the number of negative (defective) samples, and $n^- << n^+$. For under-sampling, the data set used for CNN training contains n^- negative samples, and $\hat{n}^+$ random positive samples with $\hat{n}^+ = n^-$. For over-sampling, the training data set contains n^+ positive samples, and $\hat{n}^-$ negative samples. Each collected negative sample is replicated int[n^+ / n^-] times so that $\hat{n}^- = n^+$. For class weights, the data set contains positive samples and true negative (defective) samples with respective weights $w^+ = (\frac{n^-}{n^+ + n^-})^p$ and $w^- = (\frac{n^+}{n^+ + n^-})^p$, where w^+ and w^- are respectively the weights assigned to positive and negative samples.

In the experiment, 90 true white sawmark samples and 60 black sawmark samples are used for defect synthesis, and additional 350 defect-free samples and 100 true defective samples are used for CNN testing. Each sample is of size . For the under-sampling experiment, n^- =150 and n^+ =150. It results in a recognition rate of 81.5%.

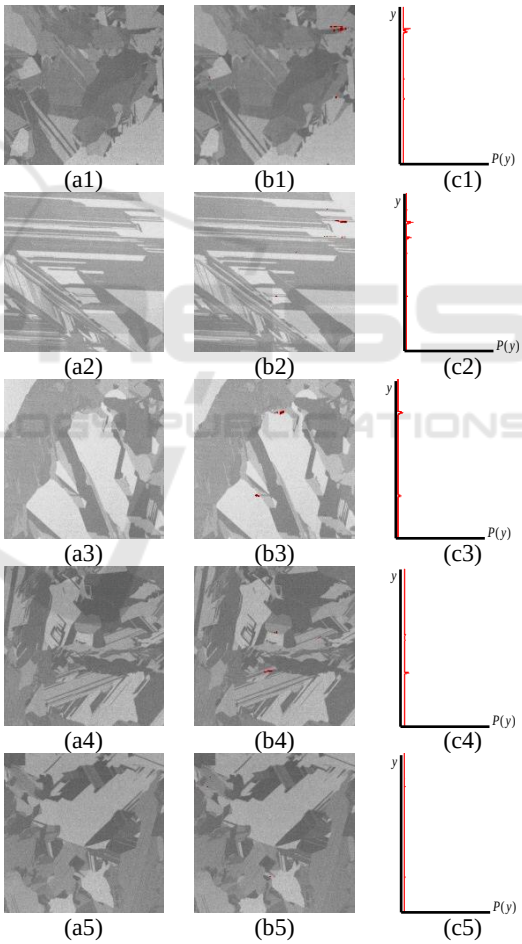

Figure 6: Detection results of defect-free solar wafer images: (a1)-(a5) faultless test samples; (b1)-(b5) suspected defect pixels (shown in red) detected by CNN; (c1)-(c5) horizontal projection profile P(y).

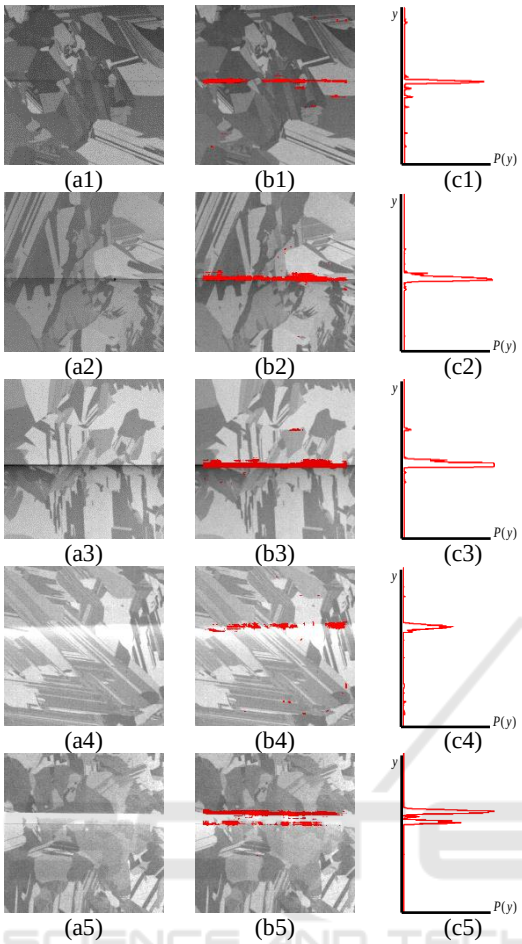

Figure 7: Detection results of defective solar wafer images: (a1)-(a5) defect samples; (b1)-(b5) detected defect pixels (shown in red) by CNN; (c1)-(c5) horizontal projection profile P(y).

Figure 8 shows the recognition rates of the over-sampling, class-weights and the proposed method with varying total number of samples used for CNN training (half of true defect-free samples and half of replicated/synthesized defective samples). It shows that the proposed method outperforms the other three comparative methods. The recognition rate of the proposed method increases as the total number of training samples is increased.

Figure 8 visually displays the detection results of the four comparative methods for two normal solar wafers and two defective solar wafers. Figures 8(a1)-(a2) are defect-free wafer images, and (a3)-(a4) are defective wafer images with respective black and white sawmarks. Figures 8(b1)-(b4), (c1)-(c4), (d1)-(d4) and (e1)-(e4) are the detection results of the CNN models trained with the proposed method, under-sampling, over-sampling and class weights. As expected, the under-sampling approach creates severe

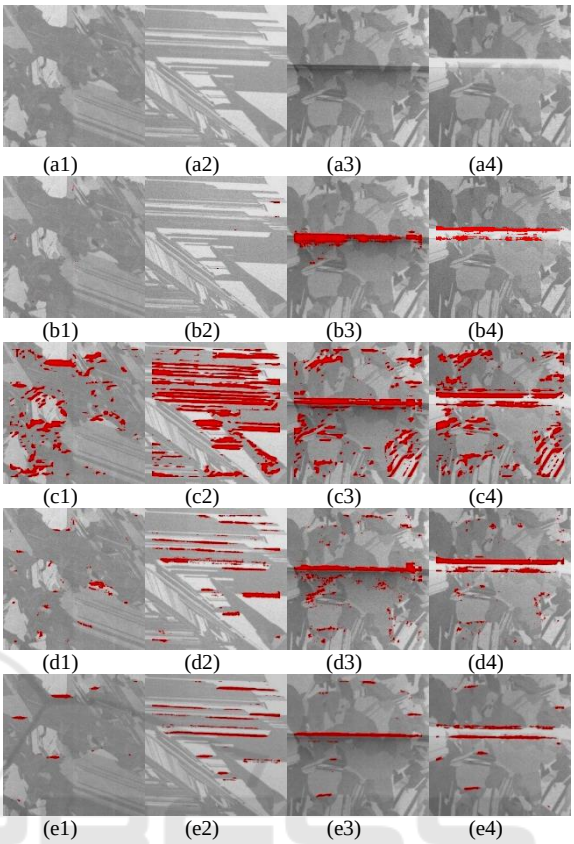

Figure 8: Comparison of defect detection of the four comparative methods for imbalanced data:(a1), (a2) defect-free; (a3)-(3) black sawmark, (a4) white sawmark; (b1)-(b4) proposed method; (c1)-(c4) undersampling; (d1)-(d4) oversampling; (e1)-(e4) class weights.

false detection in the defect-free regions. The over-sampling and class-weight approaches improve the detection capability of true defects in the solar surfaces. They generate quite a few noisy points, and severely identify the horizontal grain edges as sawmark defects. The proposed method can successfully detect white and black sawmarks with minimum noise.

4 CONCLUSIONS

The proposed paper has presented an automatic defect detection scheme to identify saw-mark defects in multicrystalline solar wafer images. The heterogeneous background of crystal grains in a solar wafer image and the saw-mark defect in a small window patch are classified by a CNN model. To overcome the shortage of defect samples in solar-wafer manufacturing and the imbalanced data

problem in CNN model training, the CycleGAN model is applied to generate a sufficiently large dataset of negative samples from a very limited number of real saw-mark image patches. Due to the indiscriminate patterns between the regular random crystal grains and the saw-mark in a small image patch, the detected saw-mark region in a full-sized solar image may not be completely detected. The postprocessing with the horizontal projection in the segmented binary image can effectively identify the presence/absence of a saw-mark in the inspection image. The preliminary experimental results indicate the proposed method can effectively detect various saw-mark defects including black line, white line and impurity in solar wafer surfaces.

The proposed method currently focuses on saw-mark detection in multicrystalline solar wafers. In the future, the use of the CycleGAN or GAN-variant models to create various defect types such as contaminants, particles and fingerprints and training the CNN model for multiple-classes classification are worthy of further investigation.

Table 1: Recognition rates with varying number of training samples for the CNN models.

number of samples	1000	2000	4000	6000	8000	10000	12000
over-sampling	88.89	91.78	91.11	92.44	90.89	93.11	90.89
class-weights (p=1/3)	90.89	92.00	92.22	90.67	90.00	90.00	90.44
proposed method	91.11	92.67	92.89	93.56	95.11	94.67	95.33

REFERENCES

Z. Fu, Y. Zhao, Y. Liu, Q. Cao, M. Chen, J. Zhang, J. Lee, 2004. “Solar cell crack inspection by image processing,” Int’l. Conf. on Business of Electronic Product Reliability and Liability, Shanghai, China, pp. 77–80.

M. A. Ordaz, G. B. Lush, 2000. “Machine vision for solar cell characterization,” Proc. of SPIE, San Diego, CA, USA, pp. 238–248.

M. Pilla, F. Galmiche, X. Maldague, 2002. “Thermographic Inspection of Cracked Solar Cells,” Proc. of SPIE, Seattle, WA, USA, pp. 699–703.

D. M. Tsai, C. C. Chang, S. M. Chao, 2010. “Micro-crack inspection in heterogeneously textured solar wafers using anisotropic diffusion,” Image and Vision Computing, vol. 28, pp. 491–501.

I. Ar, Y. S. Akgul, 2008. “A generic system for the classification of marble tiles using Gabor filters,” International Symposium on Computer and Information Sciences, Istanbul, pp. 1–6.

X. Xie, M. Mirmehdi, 2005. “Localising surface defects in random color textures using multiscale texem analysis in image eigenchannels,” IEEE Int’l. Conf. on Image Processing, Genoa, Italy, pp. III–1124–7.

X. Xie, M. Mirmehdi, 2007. “TEXEMS: Texture exemplars for defect detection on random textured surfaces,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, pp. 1454–1464.

W.-C. Li, D.-M. Tsai, 2011. “Automatic saw-mark detection in multicrystalline solar wafer images,” Solar Energy Materials and Solar Cells, vol. 95, pp. 2206-2220.

Y. LeCun, Y. Bengio, G. Hinton, 2015. “Deep learning,” Nature, vol. 521, pp. 436-444.

D. Soukup, R. Huber-Mork, 2014 . “Convolutional neural networks for steel surface defect detection from photometric stereo images, Intl. Symposium on Visual Computing, pp. 668-677.

Y. Li, W. Zhao, J. Pan, 2017. Deformable patterned fabric defect detection with Fisher criterion-based deep learning,” IEEE Trans. Automation Science and Engineering, vol. 14, pp. 1256-1264.

Y.-J. Cha, W. Choi, O. Buyukozturk, 2017. “Deep learning-based crack damage detection using convolutional neural networks,” Computer-aided Civil and Infrastructure Engineering, vol. 32, pp. 361-378.

X. Gibert, V. M. Patel, R. Chellappa, 2017. ”Deep multitask learning for railway track inspection,” IEEE Trans. Intelligent Transport. Systems, vol. 18, pp. 153-164.

J.-Y. Zhu, T. Park, P. Isola, A. Efros, 2017 . “Unpaired image-to-image translation using cycle-consistent adversarial networks,” arXiv:1703.10593v2, 5 Oct..

A. Krizhevsky, L. Sutskever, G. Hinton, 2012. “ImageNet classification with deep convolutional neural networks,” Advances in Neural Information Processing Systems 25 (NIPS).

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, 2014. “Generative adversarial nets,” Advances in Neural Information Processing Systems 27 (NIPS).

T. Zhou, P. Krahenbuhl, M. Aubry, Q. Huang, A. A. Efros, 2016. “Learning dense correspondence via 3d-guided cycle consistency,” CVPR, pp. 117-126.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. 2002. SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321-357.

Yen, S. J., and Lee, Y. S. 2009. Cluster-based under-sampling approaches for imbalanced data distributions. Expert Systems with Applications, 36, 5718-5727.