Vision based wafer states detection in front opening unified pod load-port system

*Qiang* Zhang[1,*], *Xueying* Sun[1], and *Mingmin* Liu[2]

[1]School of Electronic Information, Jiangsu University of Science and Technology, No. 666 Changhui Road, Zhenjiang, 212003 China
[2]SIASUN Robot & Automation CO., Ltd, No.33 Quanyun Road, Shenyang 110169, China

Abstract. In modern integrated circuit manufacturing processes, wafers are always transported from one procedure to another. To reduce the risk of dust, Front Opening Unified Pod (FOUP) load-port system is always adopted. Misplaced wafers should be detected before transported. Traditional methods always fail to detect wafer states correctly. To improve detection accuracy, this paper proposed a vision based method. Wafer overlap and malposition detection approach based on modified YOLO-V3 algorithm was suggested. Experiment results shows superiority of the proposed approach.

1 Introduction

In integrated circuit manufacturing, silicon chips are produced through multiple production processes. Wafer handling system plays a very important role in the factory for transporting wafers from one work flow to another. FOUP load-port system, which can be seen in Figure 1, is an equipment for recovering or placing silicon wafers from the processing unit through wafer box. In the FOUP load-port system, vacuum robot is used as loading and unloading mechanism.
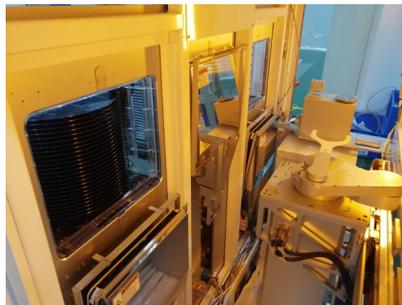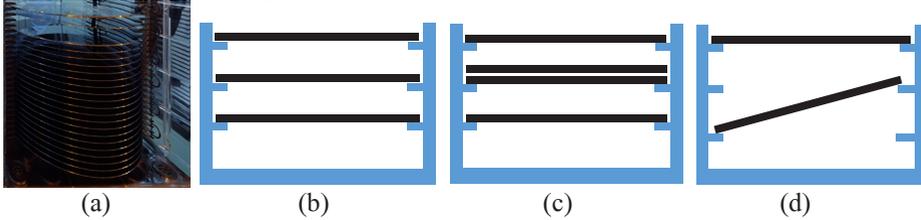

Fig. 1. FOUP load-port system.

In some cases, wafers in the FOUP have the possibility of overlap or malposition. This would cause subsequent operations to fail. Overlap and malposition states can be seen in

---

* Corresponding author: qzhang@just.edu.cn

Figure 2. In order to correct these mistakes, wafer overlap or malposition detection procedures should take effect. However, traditional detection methods, for example opposite-type photoelectric switch based methods, cannot meet the requirements of high-precision detection. As a result, FOUP load-port system is hardly competent in the integrated circuit processing.

    (a)   (b)   (c)   (d)

Fig. 2. Silicon wafer states: (a)FOUP and wafers; (b)correct; (c)overlap; (d)malposition.

In our work, we designed a vision based detection system to improve wafer overlap and malposition detection accuracy in FOUP load-port applications. Firstly, we choose monocular camera as the sensor. In order to improve system's robustness, we carefully adjusted the position and the angle of view of the camera. Secondly, we proposed a wafer overlap and malposition detection approach based on improved YOLO-V3 algorithm. According to experiment results, the proposed method can provide high detection accuracy and efficiency.

The rest of the paper is organized as follows. Section 2 introduces methods for object detection. Section 3 discussed details of the proposed approach for silicon wafer detection. Section 4 describes performance of the detection system and analyses experiment results. The last part, Section 5, gives the conclusion of the paper.

2 Related works

In recent few years, some effective methods have been proposed to detect objects in images. Traditional detection methods like keypoint feature matching based methods[1–3], shape matching based methods[4,5], Hough transform based methods[6], machine learning based methods[7,8] were developed to recognize or detect objects in images.

From year 2014, deep learning based methods were developed rapidly. Many deep learning architectures like Fast R-CNN[9], SSD[10], YOLO[11] have been developed to revolutionize object detection accuracy. As a regression based algorithm, YOLO and the subsequent improved algorithm get more and more attentions. This paper introduced a method inspired by YOLO-V3[12] architecture, and the approach was successfully applied in wafer states detection.

3 Robust wafer overlap and malposition detection approach

Original images have lots of background information. These background information has no contributions to the detection results. To improve detection robustness, only regions containing wafers are reserved. Then, projective transformation and downsampling are performed to obtain 416×416 size images. And the results are feed into our designed detection neural network.
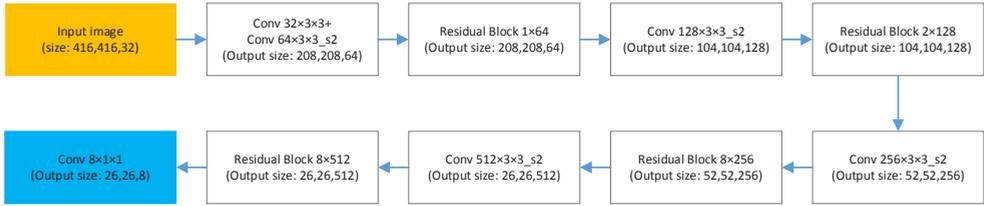
Fig. 3. The proposed neural network

The designed approach is inspired by YOLO-V3 architecture. The difference between YOLO-V3 and ours are reflected in two aspects. The first is that we simplified the original YOLO-V3 network. The second is that the output results are truncated to meet the detection needs. The structure of the proposed neural network can be seen in Figure 3.

In the wafer states detection system, we divide waters' states into three classes: correct, overlap and malposition. To address the detection problem, each detection result is represented as a $26\times26\times8$-dimentional tensor. The result indicates that there are 26×26 detection result and each result is a 8-dimentional vector, which can be written as $\{t_x, t_y, t_w, t_h, p, C_{correct}, C_{overlap}, C_{malposition}\}$. In the 8-dimentional vector, $\{t_x, t_y, t_w, t_h\}$ indicates the predicted bounding box. We first give a reference bounding box to help the network to obtain accurate detection results. Reference bounding box is chosen by clustering labeled data. For the reason that the size of the reflective edges of wafers are similar, only one reference bounding box is chosen. In the detection result, t_x and t_y are offsets between predicted center and reference bonding box center. t_w and t_h are width and height scale ratios between predicted box and reference bounding box. p is the confidence score of the predicted bounding box. $C_{correct}$, $C_{overlap}$ and $C_{malposition}$ are correct, overlap and malposition class probabilities.

4 Experiment and results analysis

4.1 Methodology

In our experiment, PyTorch open source library was employed and CUDA computation acceleration toolkit was introduced to training and deploy procedures. The software and hardware configurations can be seen in Table 1.

**Table 1.** Hardware configuration in the experiment.

Hardware	Parameter
CPU	Intel i7-10875H 2.3GHz
Computer Memory Size	16GB
Graphics card	RTX2070 8GB
Operating system	Ubuntu 18.04
Deep learning library	PyTorch 1.3.0

We collected 500 images from the industrial site. LabelImg annotation tool was employed to label the acquired image data. Image cropping and rotation were also executed to extend the dataset. Image number of the dataset was extended to over 2000 and the number of wafers' labels exceeded 35000. In the labeled dataset, about 50% were correctly placed wafers, overlap and malposition classes take about 25% respectively. In the experiment, the training dataset was divided to training dataset and validation dataset randomly. The training dataset occupies 90% of the whole dataset and the validation dataset takes 10%.

During training process, parameters of the last layer were initialized randomly. Other parameters were initialized by pre-trained original YOLO-V3 neural network. The initial learning rate was set to 0.001 with attenuation factor of 0.1 for each 5000 iterations. The maximum iteration number was set to 20000. In order to verify the improved method, we did a comparative experiment using original YOLO-V3 algorithm.

4.2 Experiment results

Training loss with iteration can be seen in Figure 4(a). After 20000 iterations, training stage was stopped. PR curve can be seen in Figure 4(b).

(a)Loss with iteration curve   (b)PR curve
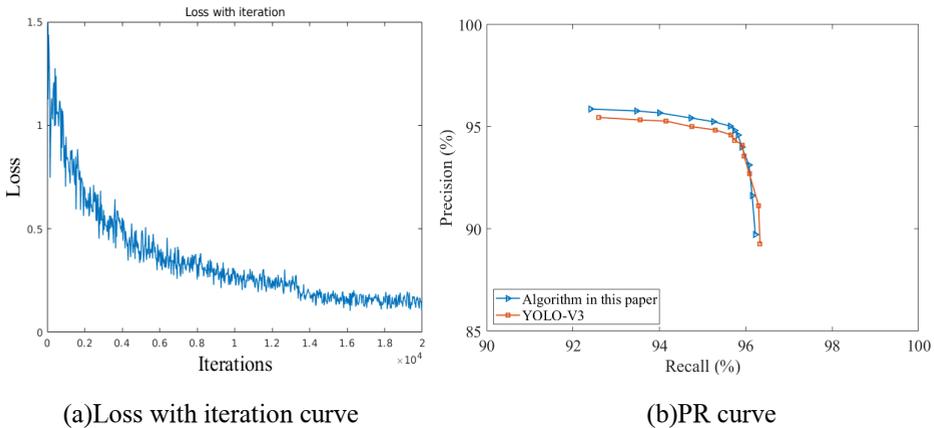
Fig. 4. Training loss curve and PR curve: (a)Loss with iteration curve; (b) PR curve

According to the experiment results in Figure 4(b), the proposed method has a better performance in precision. Precision is improved for the reason that the designed architecture output is more reasonable for wafer state detection. And with the help of GPU, the detection efficiency can reach to 35fps. This detection accuracy and speed can meet the needs of industrial field.

5 Conclusions

Wafer state detection is critical in integrated circuit manufacturing. This paper introduced a monocular camera based wafer state detection system. In the system, modified YOLO-V3 algorithm was employed to detect wafer states. For the reason that the imaging results are relatively fixed, we optimized the YOLO-V3 architecture according to the specific application scenarios. With the proposed method, the system can achieve good performance in industrial applications.

References

1. D. G. Lowe. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision, **60**(2): 91–110 (2004)

2. H. BAY, T. Tuytelaars, L. V. Gool. SURF: Speeded Up Robust Features[J]. Computer Vision & Image Understanding, **110**(3): 404–417 (2006)

3. E. Rosten, R. Porter, T. Drummond. Faster and better: a machine learning approach to corner detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(1): 105–119 (2008)

4. S. Hinterstoisser, V. Lepetit, S. Ilic, P. Fua, N. Navab. Dominant Orientation Templates for Real-Time Detection of Texture-Less Objects[C]. IEEE Conference on Computer Vision and Pattern Recognition (2010)

5. S. Hinterstoisser, C. Cagniart, S. Ilic, P. Sturm, N. Navab, P. Fua, V. Lepetit. Gradient Response Maps for Real-Time Detection of Textureless Objects[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, **34**(5): 876–888 (2012)

6. Y. Ito, K. Ogawa, K. Nakano. Fast Ellipse Detection Algorithm Using Hough Transform on the GPU[C]. International Conference on Networking & Computing (2011)

7. X. Ren, D. Ramanan. Histograms of Sparse Codes for Object Detection[C]. IEEE Conference on Computer Vision & Pattern Recognition (2013)

8. E. Gadeski, H. O. Fard, H. L. Borgne. GPU deformable part model for object recognition[J]. Journal of Real-Time Image Processing, **14**(2): 1–13 (2014)

9. R. B. Girshick. Fast R-CNN[C]. 2015 IEEE International Conference on Computer Vision (ICCV), 1440–1448 (2015)

10. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, A. C. Berg. Ssd: Single shot multibox detector[C]. European conference on computer vision, 21–37 (2016)

11. J. Redmon, S. Divvala, R. Girshick, A. Farhadi. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779–788 (2016)

12. J. Redmon, A. Farhadi. YOLOv3: An Incremental Improvement[J]. arXiv:1804.02767 [cs], (2018)