IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 7, 2021, accepted November 2, 2021, date of publication November 8, 2021, date of current version November 12, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3125735

3D Model Registration-Based Batch Wafer-ID Recognition Algorithm

FANG CAO[1], ZENGGUO TIAN[1,3], BAOZHU JIANG[1], HONGSHUAI ZHANG[2], HENG CHEN[1], AND XUGUANG ZHU[3]

[1]School of Physics and Microelectronics, Zhengzhou University, Zhengzhou 450001, China
[2]School of Mechanical and Power Engineering, Zhengzhou University, Zhengzhou 450001, China
[3]Luoyang Dejing Intelligent Technology Company Ltd., Luoyang 471032, China

Corresponding authors: Zengguo Tian (tianzg@zzu.edu.cn) and Hongshuai Zhang (hszhang@zzu.edu.cn)

This work was supported in part by the Special Fund for Advanced Manufacturing Development of He'nan Province of China.

ABSTRACT Wafer identification (ID) is a serial number printed on the surface of wafer, which is used for indexing production process data in manufacture execution system. The automatic recognition of wafer ID is helpful to improve the level of automatic production. However, the existing equipment and methods mainly focus on single wafer-ID recognition, which require wafers to be taken out and placed on a specific platform, resulting in low efficiency. In this paper, we present a batch wafer-ID recognition method based on machine vision, including a specific designed image-acquisition system and recognition algorithms. Based on the priori information, we formulate a 3D model for the cassette and wafers to be registered with the features extracted from the image. Combined with image-processing techniques, the pose of wafers in the cassette is estimated to undistort the perspective deformation of wafer-ID characters, such that we can exploit a classic lightweight convolution neural network for character recognition. The proposed system can capture and recognize the whole image of a cassette of wafers at once, which does not need to take out the wafers from the cassette and avoids the risk of contamination. Extensive experiments were conducted to evaluate the performance of our proposed techniques by collecting the data set of wafer-ID images. The results show that our proposed 3D model-based rectification method can correct the character deformation effectively and enable the lightweight classifier to achieve high speed and high accuracy for batch wafer-ID recognition.

INDEX TERMS Optical character recognition, wafer identification, machine learning, model registration.

I. INTRODUCTION

With the rapid development of large-scale integrated circuit, silicon wafer-based semiconductor devices have been widely used [1]. To ensure the quality of products, tracing the product quality is crucial [2]. In semiconductor industries, laser marking is generally printed on the wafer surface as a unique identification (ID). Subsequently, the wafer IDs are recognized in the subsequent processes and recorded in the manufacture execution system. Therefore, a large number of IDs need to be recognized during silicon wafer production. The traditional method of recognizing the wafer ID includes its manual reading and inputting into the computer, which is an inefficient and labor-intensive way that hinders the development of automatic production.

Optical character recognition (OCR) is a key technology to realize the rapid automatic input and is widely used for many specific applications, such as the recognition of forms [3], certificates [4], and license plates [5]. Existing research on OCR seldom focuses on wafer-ID recognition [6], [7] and most of them tend to recognize single wafer-ID by the piece. Several wafer ID readers, such as Sorter-OCR [8], WID-110 [9], and ID-Scope [10], are available on the market. Owing to the tiny characters and shallow marks of the ID on the wafer, these devices depend on special optical structure to obtain the clear image of each wafer ID, which removes the difficulty in recognition. Sorter-OCR is a typical product for wafer-ID recognition developed by R2D Automation Co., Ltd. This device adopts a mechanical way to take out and place back the wafer to capture the image of a wafer ID, which takes up to 2 minutes to recognize all the wafer IDs in a cassette. As for WID-110 and ID-Scope, novel mechanical structures are designed to push the wafer for piece-by-piece

The associate editor coordinating the review of this manuscript and approving it for publication was Amin Zehtabian.

recognition without taking the wafer out of the cassette, thereby accelerating image acquisition. However, recognizing all wafer-IDs in a cassette still takes 35 seconds, which cannot meet the needs of the mass production of silicon wafer. In addition, the mechanical operation of wafers introduces the probability for mechanical damage and additional risk of contamination. Therefore, a new image acquisition and recognition method needs to be proposed for batch wafer-IDs recognition.

In this work, we aim to design an OCR system for wafer-ID recognition to improve the automation level of wafer fabrication. We present a batch wafer-ID recognition system to realize high recognition efficiency and accuracy. In this system, a specific image acquisition scheme is designed for the recognition of multiple wafer-IDs, with a specific base for wafer cassette, linear light sources, and a high-resolution camera. Images are taken at an inclined angle, and all the wafer-IDs in a cassette can be captured in an image at one time without any contact operation on the wafer. The wafer cassette only needs to be placed on the base of the cassette without damaging and contaminating the wafers. Then, OCR algorithms, such as deep learning-based methods, can be used for batch wafer-ID recognition. However, the perspective deformation of characters and the long number of wafer-ID make the training of the neural network to achieve high accuracy difficult. To solve this problem, based on the proposed image acquisition scheme, we formulate the priori 3D model and incorporate the line features detected in the image to estimate the pose of each wafer in a cassette, which is the so-called 3D model-based registration. Subsequently, the estimated poses are used to accurately correct the character undistortion with perspective transformation, which alleviates the pressure of character recognition. Finally, a shallow convolution neural network (CNN) is trained to realize fast and accurate wafer ID recognition using these undistorted images. In summary, the main contributions of the presented work are detailed as follows:

- We propose an image acquisition scheme that can capture all the wafer-IDs in a cassette simultaneously. This scheme does not need any contact operation with wafers, thereby accelerating the recognition and avoiding the damage and contamination of wafers.
- We present a 3D model registration-based rectification method to correct the deformation of character images, which further enables a light weight CNN to achieve high accuracy.
- We make datasets to evaluate our algorithms by capturing many pictures of wafer-IDs using the proposed image acquisition scheme. Extensive experiments have been conducted and have shown the effectiveness and advantages of our method.

II. RELATED WORK

With the rapid development of computer vision and artificial intelligence, the efficiency and accuracy of OCR techniques have been improved greatly. Traditional OCR is based on image processing and statistical machine learning. A typical traditional OCR framework consists of text region location, text image undistortion, character segmentation, and character recognition. Text image deformation is generally caused by the change in the view point of a camera. Some geometric corners of the image of paper form or license plate can be used to calculate perspective transformation to deskew the character image. Then, the undistorted image can be classified with machine learning-based methods [11], such as AdaBoost and SVM. However, no corners can be used for wafer-ID image correction, such that various character deformation drags down the precision of traditional classification algorithm. To address the aforementioned problems, a wafer-ID segmentation and recognition method for low quality image is proposed in [6]. This method uses the combined relationship of multiple templates to match various features of the detection image. In [7], a wafer-ID scheme based on an asterisk-shape filter is proposed to examine the serious influence of uneven brightness on wafer-ID recognition. Nevertheless, these methods are designed to recognize single wafer-ID by the piece and suffer from low efficiency.

The development of deep neural networks has boosted character recognition. Given abundant training on data set, deep neural networks can learn the model to detect text region and recognition characters. Some representative methods include CTPN [12] and DMP-Net [13] for text detection, as well as CRNN [14] and DTRN [15] for text recognition. To enable the network to recognize deformed characters, spatial transform network (STN) [16] is proposed. STN is an unsupervised learning method that can correct the distortion of characters by self-learning the back propagation gradient. However, the improvement of character recognition is not apparent. Based on STN, ASTER [17] is proposed to make further step by exploiting two rows of control points to roughly fit the character boundary. These two rows of control points are located on the upper and lower boundaries of the character region and correspond one by one. The fitting area of the control points can be corrected with thin plate-spline interpolation. Although it can alleviate the influence of character deformation within a small range near the target text better, it cannot completely recover the deformation, which still leads to the large-scale training of recognition network for high accuracy.

Extensive research on text recognition in complex natural environments has been witnessed in recent years [18]–[21]. The text perceptron method [18] combines text detection and recognition into an end-to-end network, which proposes a segmentation-based method to detect the contour points of the text for correcting the text region. Li *et al.* [19] exploited parallel detection and recognition methods to extract scene text in an end-to-end manner. They established a weak connection between the detection module and the identification module to guide the update of model parameters, which provided a method to balance the detection and identification of the impact on system performance. Geometric normalization

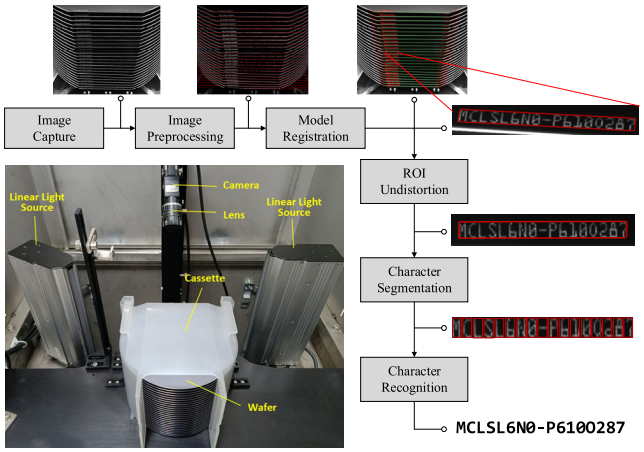


FIGURE 1. Outline of the proposed wafer-ID recognition system.

networks [20] are proposed for deformed text detection with irregular shape and scale. By transforming the feature map of the text image geometrically, the text boxes with large differences in geometric distribution can be normalized to a limited range, which improves text detection in natural scene. Liu *et al.* presented an adaptive Bezier curve network, which involve a novel Bezier alignment layer to extract the convolution features of arbitrary shape text, to detect scene text in real time [21]. With the parameterized Bezier curve, the method can adaptively fit the text shape with very low computational overhead. Although the existing research on natural scene text has overcome arbitrary direction and shape to a certain extent, the recognition accuracy is still significantly lower than that of straight-line text. The direct application of these methods to the actual scene does not provide satisfactory results. Therefore, deformed text recognition still faces major problems that need to be solved.

III. SYSTEM ARCHITECTURE

Our wafer-ID recognition system is proposed to capture a clear image of batch wafers in a cassette and recognize all the wafer-IDs at one time. The system hardware consists of light sources, lens, high-resolution camera, and personal computer. In this work, we focus on recognizing the ID of a six-inch wafer. According to the national standard of wafer marking [22], the ID characters are printed near the edge line with laser marking in the form of dot matrix. The size of the character area on wafer is 1.624 mm $\times$ 0.812 mm, and the depth of the marks is 0.2 mm. Considering the tiny characters, shallow marking, and the high reflection characteristic of wafer, a specific light-source scheme with a proper view direction must be selected to ensure the quality of the image captured from the camera. The outline of our system is shown in Figure 1. We design a specific image-acquisition method to obtain a clear image with linear-light sources placed at both sides of the cassette and a high-resolution camera posed at a top view angle. The wafers being recognized are placed in a cassette that contains 25 pieces of wafers. In addition, we design a specific base for a cassette to ensure the cassette is placed in the same position in each image. The light sources illuminate wafers within a small angle, thereby making the wafer-ID characters and wafer edge particularly clear in the image. Hence, we can clearly obtain all the wafer IDs within an image.

Once an image of wafer IDs in a cassette is received, it is inputted into our proposed recognition pipeline for further processing. Our wafer-ID recognition pipeline contains image preprocessing, character undistortion, and character recognition (Figure 1). The outline of the processing steps is briefly demonstrated as follows:

First, we receive the image from the high-resolution camera and preprocess it with median filter, adaptive binarization, and line detection. We use median filter and adaptive binarization to remove noise and to overcome local brightness changes, respectively. In the step of line detection, the edge line of each wafer in the binary image is identified, and the parametric equations of each line is calculated for subsequent 3D model-based undistortion. As shown in Figure 1, we repaint the detected lines on the binarized image to visualize the preprocessing result.

Second, based on the 3D model of the wafers in a cassette, the projection of edge lines from the 3D model to the image are aligned with the actually detected lines in the image, so as to estimate the orientation of each wafer. With the estimated pose of each wafer, the actual character region of each wafer, which is a quadrilateral area in the image, can be calculated. Next, we use perspective transformation to correct the quadrilateral region into a rectangle region, so as to realize the character undistortion.

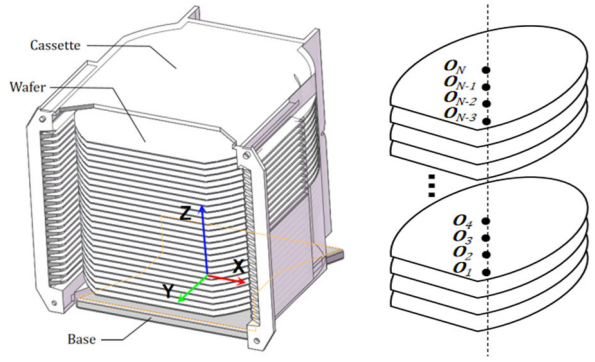
Finally, the rectangle character region is segmented using a stable region analysis method to separate each single character into an image. Then, we extract the image features of the segmented characters and input them into the trained classifier for the character recognition of wafer IDs. After classifying all the images, the system output the recognition results.

IV. METHOD

A. 3D MODEL FORMULATION

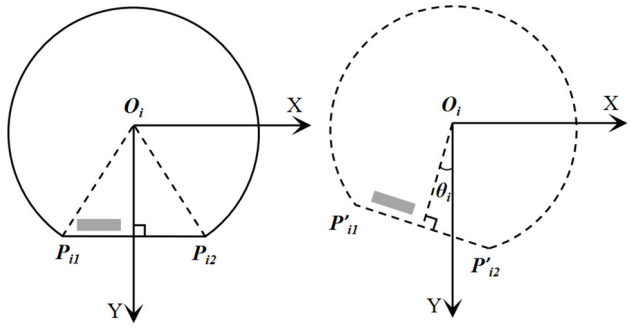
The character area exhibits different degrees of perspective deformation because the camera is inclined to take pictures and the pose of each wafer is different. If the characters are not corrected, then the deformed image may lead to the failure of segmentation and degrade the accuracy of character recognition. Inspired by the perspective transformation-based undistortion method for license-plate image, we propose a 3D model-based registration method for the perspective-transformation undistortion of wafer-ID images.

First, we construct the 3D model of the wafers in a cassette in the world coordinate (Figure 2). Herein, we assume these wafers are concentric because the pre-sequence production process has arranged the wafers in a cassette orderly. We define the center of the upper surface of the cassette



a) Wafers in a cassette b) Wafers and their centers

FIGURE 2. 3D model of the wafers in a cassette.



a) Wafer without rotation b) Wafer with rotation

FIGURE 3. Top view of a wafer with/without rotation. The gray rectangle represents the character region of wafer ID.

base as the world coordinate origin $\boldsymbol{O}_W$. The center of each wafer is denoted as $\boldsymbol{O}_i = [0, 0, h_i]^T$, $i \in \{1, \ldots, N\}$. According to the 3D model, the height h_i can be calculated as follows:

$$h_i = h_0 + (i-1) \cdot \Delta h. \tag{1}$$

Then, we denote R as the radius of the wafer and L as the length of the edge line $\boldsymbol{P}_{i1}\boldsymbol{P}_{i2}$ on the wafer (Figure 3a). The two end points of the edge line can be represented as:

$$\begin{cases} \boldsymbol{P}_{i1} = \left[\frac{1}{2}L, \sqrt{R^2 - \frac{1}{4}L^2}, h_i\right]^T \\ \boldsymbol{P}_{i2} = \left[-\frac{1}{2}L, \sqrt{R^2 - \frac{1}{4}L^2}, h_i\right]^T. \end{cases} \tag{2}$$

We assume that each wafer in a cassette has only one degree of freedom, which corresponds to the rotation angle of yaw θ_i (Figure 3b). Then, Formula (2) can be further formulated as:

$$\begin{cases} \boldsymbol{P}'_{i1} = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 \\ \sin\theta_i & \cos\theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{P}_{i1} \\ \boldsymbol{P}'_{i2} = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 \\ \sin\theta_i & \cos\theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{P}_{i2}. \end{cases} \tag{3}$$

B. 3D MODEL REGISRATION

We calibrate the camera to obtain the intrinsic matrix $\boldsymbol{K}$ and the extrinsic matrix $(\boldsymbol{R}_{CW}, \boldsymbol{t}_{CW})$ with Zhang [23] method to obtain the transformation between image coordinate and world coordinate. Then, a point $\boldsymbol{P}_W = [X, Y, Z]^T$ in the world coordinate can be denoted as a point $\boldsymbol{P}_C = [X_C, Y_C, Z_C]^T$ in the camera coordinate using the following formula:

$$\boldsymbol{P}_C = \boldsymbol{R}_{CW}\boldsymbol{P}_W + \boldsymbol{t}_{CW}. \tag{4}$$

The point $\boldsymbol{P}_W$ in the world coordinate can be projected into the image coordinate with:

$$\boldsymbol{p} = \frac{1}{\boldsymbol{Z_C}}\boldsymbol{K}\boldsymbol{P_C} = \frac{1}{\boldsymbol{Z_C}}\boldsymbol{K}\left(\boldsymbol{R_{CW}}\boldsymbol{P_W} + \boldsymbol{t_{CW}}\right). \tag{5}$$

Therefore, the 3D end points of each wafer are projected to the image plane to obtain the coordinates $\boldsymbol{p}_{i1} = \left[u_{p_{i1}}, v_{p_{i2}}, 1\right]^T$ and $\boldsymbol{p}_{i2} = \left[u_{p_{i2}}, v_{p_{i2}}, 1\right]^T$:

$$\begin{cases} \boldsymbol{p}_{i1} = \frac{1}{Z_C}\boldsymbol{K}\left(\boldsymbol{R}_{CW}\boldsymbol{P}'_{i1} + \boldsymbol{t}_{CW}\right) \\ \boldsymbol{p}_{i2} = \frac{1}{Z_C}\boldsymbol{K}(\boldsymbol{R}_{CW}\boldsymbol{P}'_{i2} + \boldsymbol{t}_{CW}). \end{cases} \tag{6}$$

The slope k_i of a straight line $\boldsymbol{l}_i$ can be determined by two projected points, namely, $\boldsymbol{p}_{i1}$ and $\boldsymbol{p}_{i2}$:

$$k_i = (u_{p_{i1}} - u_{p_{i2}})/(v_{p_{i1}} - v_{p_{i2}}). \tag{7}$$

We can further match the projected end points with the detected lines $\boldsymbol{l}'_i$ in the image. Then, we define the cost function to search for the rotation angle θ_i of each wafer:

$$\theta_i^* = argabs\left(k'_i - k_i\right), \quad \theta_i \in [-5, 5], \tag{8}$$

where k'_i denotes the slope of the detected line $\boldsymbol{l}'_i$, and the projected line and the detected line are associated by the height h_i of the i-th wafer in the 3D model. Herein, we set the searching scope within $\pm 5°$ because the pre-sequence production process has neatly arranged the wafers in the cassette with minor rotation.

C. PERSPECTIVE TRANSFORMATION

According to the character regions of the wafer ID specified in the standard [22], we can calculate the four corners of the character area on each wafer as follows:

$$\begin{cases} \boldsymbol{Q}_{i1} = [x_{i1}, y_{i1}, h_i]^T \\ \boldsymbol{Q}_{i2} = [x_{i2}, y_{i2}, h_i]^T \\ \boldsymbol{Q}_{i3} = [x_{i3}, y_{i3}, h_i]^T \\ \boldsymbol{Q}_{i4} = [x_{i4}, y_{i4}, h_i]^T. \end{cases} \tag{9}$$

Considering the calculated wafer rotation angle θ_i, Formula (9) can be further represented as:

$$\begin{cases} \boldsymbol{Q}'_{i1} = \boldsymbol{R}_\theta \boldsymbol{Q}_{i1} \\ \boldsymbol{Q}'_{i2} = \boldsymbol{R}_\theta \boldsymbol{Q}_{i2} \\ \boldsymbol{Q}'_{i3} = \boldsymbol{R}_\theta \boldsymbol{Q}_{i3} \\ \boldsymbol{Q}'_{i4} = \boldsymbol{R}_\theta \boldsymbol{Q}_{i4}, \end{cases} \tag{10}$$

$$\boldsymbol{R}_\theta = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 \\ \sin\theta_i & \cos\theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{11}$$

Furthermore, we use the external extrinsic and intrinsic matrices of the camera to project the four corner points of the character region on the 3D model to obtain the coordinates of the four corners on the image:

$$\begin{cases} \boldsymbol{q}_{i1} = \dfrac{1}{Z_C}\boldsymbol{K}\left(\boldsymbol{R}_{CW}\boldsymbol{Q}'_{i1} + \boldsymbol{t}_{CW}\right) \\ \boldsymbol{q}_{i2} = \dfrac{1}{Z_C}\boldsymbol{K}\left(\boldsymbol{R}_{CW}\boldsymbol{Q}'_{i2} + \boldsymbol{t}_{CW}\right) \\ \boldsymbol{q}_{i3} = \dfrac{1}{Z_C}\boldsymbol{K}\left(\boldsymbol{R}_{CW}\boldsymbol{Q}'_{i3} + \boldsymbol{t}_{CW}\right) \\ \boldsymbol{q}_{i4} = \dfrac{1}{Z_C}\boldsymbol{K}\left(\boldsymbol{R}_{CW}\boldsymbol{Q}'_{i4} + \boldsymbol{t}_{CW}\right). \end{cases} \tag{12}$$

The four corners, namely, $\boldsymbol{q}_{i1}, \boldsymbol{q}_{i2}, \boldsymbol{q}_{i3}$, and $\boldsymbol{q}_{i4}$, form a quadrangular area that bounds the characters on the image. Then, we use perspective transformation to correct the quadrilateral region into a rectangle region, so as to realize the character undistortion.

The principle of perspective transformation [24] is to project the image plane to another viewing plane. For a point $\boldsymbol{p} = [u, v, 1]^T$ on the source surface and a point $\boldsymbol{p}' = \left[u', v', 1\right]^T$ on the target surface, the perspective transformation between these two points can be formulated as follows:

$$\boldsymbol{p}' = \boldsymbol{T} \cdot \boldsymbol{p}, \tag{13}$$

where $\boldsymbol{T}$ denotes the perspective transformation matrix, which can be described as a 3×3 matrix. Considering that both the transformed points are located on a 2D plane, eight unknown elements can be found in the matrix $\boldsymbol{T}$. Then, Formula (13) can be further described as follows:

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \tag{14}$$

Given four pairs of mapping points, we can solve the perspective transformation matrix $\boldsymbol{T}$. Here, we use the four aforementioned corners of the quadrangular region as source points and four vertices of a predefined rectangle as target points. Using these four pairs of mapping points, we can calculate perspective transformation matrix T using Formula (14). Finally, we can use the perspective transformation to correct each pixel in the quadrangular region to form an undistorted image (Figure 4). After perspective transformation correction, the characters become regular without changes in tilt, rotation, and scaling, which can facilitate subsequent segmentation and recognition.

FIGURE 4. Perspective deformation-undistortion result.

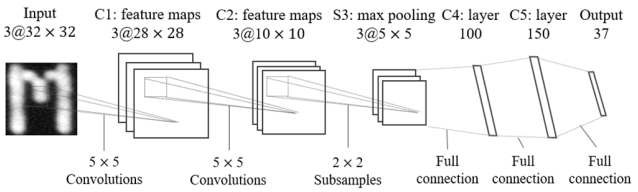


FIGURE 5. Architecture of the three-layer CNN.

D. CHARACTER RECOGNITION

1) CHARACTER SEGMENTATION

To overcome the impact of local illumination changes on segmentation, we exploit maximum stable extreme region (MSER) [25], which is immune to the image illumination change, to extract character regions. By adjusting the threshold for binarization dynamically, the region with minimal changes, which is called a MSER, can be extracted as follows:

$$\delta(i) = \left|C_{i+\Delta} - C_{i-\Delta}\right| / C_i, \tag{15}$$

where C_i represents the connected area with the binary threshold i, Δ denotes the small change in the threshold, and $\delta(i)$ corresponds to the change in the area C_i. If $\delta(i)$ is less than the given threshold δ_{th}, then the region C_i can be determined as a MSER. Meanwhile, the gray level within the detected MSER is less than the boundary, such that the white region with black background cannot be detected. To solve this problem, we need to inverse the gray level of the image and recalculate the MSER. Then, we can obtain the final result by merging the two results using the logical operator AND.

With the extracted MSERs, we can further find their 2D bounding boxes. However, overlapping bounding boxes still exist because of the broken characters on the binary image. Then, we use non-maximum suppression [26] to merge adjacent bounding boxes by calculating the intersection union ratio (IoU) of two adjacent bounding boxes as follows:

$$IoU(A, B) = (A \cap B) / (A \cup B). \tag{16}$$

If the IoU of two bounding boxes is greater than the given threshold δ_{IoU}, then the two bounding boxes should be combined. Hence, the redundant bounding boxes can be filtered out, and the single character image can be obtained.

2) CHARACTER RECOGNITION

With the step of character segmentation, the recognition difficulty is reduced. To achieve high-accuracy and high-speed recognition, we use a light-weight CNN architecture [27] for single character recognition. CNN is a kind of improved neural network structure with strong ability of feature extraction and classification, which has been widely

used in classification tasks. Typically, a CNN contains a series of convolution, pooling, and fully connected layers. With the sliding convolution operation of multi-layers, the network can learn high-level features and become a highly accurate classifier. In this paper, the classical framework of a three-layer CNN is exploited for recognition model training. As shown in Figure 5, the network works on the input image downsampled to 32×32. The first two layers are convolutional layers with kernel size 5×5. The second layer is followed by a 2×2 max-pooling layer. After the third layer, two fully-connected layers with 100 and 150 units can be observed. Finally, the output layer uses SoftMax for nonlinear fitting to obtain multi-classification result.

V. EXPERIMENT RESULTS AND DISCUSSION

A. DATASET AND IMPLEMENTATION DETAILS

We conduct extensive experiments to verify the effectiveness of our model and compare its performance with other state-of-the-art methods, including CRNN [14] and ASTER [17]. To evaluate the performance of the perspective text, we capture many images of wafer IDs using our proposed image-acquisition scheme. Accordingly, we collect 42 images of 6-inch wafers in 15 cassettes (Figure 6a). In each image, the rotation of the wafer is randomly placed within 5°. Then, we segment the sub-image of each wafer-ID region and obtain 1050 text lines. As the aforementioned methods in Section 2A-2C, we calculate the region of each wafer ID in the image based on the 3D model registration. Then, based on perspective transformation, undistorted wafer-ID images are generated. In this way, all the collected images are cropped into single wafer-ID images to build distorted and an undistorted datasets for the comparative experiments (Figures 6d–e).

We implement the proposed method using Python, OpenCV, and TensorFlow. Experiments are carried out on a laptop with a 3.4 GHz Intel(R) Core(TM) I5 8265U CPU, 16GB RAM, and an NVIDIA(R) GeForce(R) MX250 GPU with 4 GB memory. Our optical system consists of a high-resolution camera with 5472×3684 resolution, a lens with 5 mm focal length, and two white linear-light sources with 250 mm length. Owing to the simple network architecture, our CNN can be trained and tested with only CPU, whereas the evaluation of CRNN and ASTER needs to be run on GPU. In this paper, we train the CNN model using an Adam optimizer [28] by setting the learning rate $\rho = 0.9$. The thresholds of the proposed segmentation method are set as $\delta_{th} = 0.2$ and $\delta_{IoU} = 0.3$. Except for setting the size of the input image to 60×600, we use the default parameters and training methods mentioned in CRNN [14] and ASTER [17]. To train ASTER, the number of control points for ASTER is set to 20, and ADADELTA optimizer [29] with the classic learning rate schedule is adopted. The CRNN networks are also trained with ADADELTA while setting the learning rate to 0.9. In addition, the k-fold cross-validation method is exploited to enable the use of each sample data as training and test data, thereby avoiding the state of over- or under- learning effectively. Herein, we take $k = 5$ (*i.e.*, fivefold cross validation).

B. EVALUATION ON RECTIFICATION METHOD

To analyze the effect of our proposed rectification algorithm, we randomly sample 150 wafer-ID images and segment each wafer-Id image into sub-images with single character using the segmentation algorithm mentioned in Section 2D. Then we calculate the tilt angle of the text line and the clearance between two neighboring characters before and after undistortion separately. As shown in Figure 7, we calculate the angle between the text line and the horizontal, as well as the pixel number between the adjacent bounding boxes of two single-character. If two bounding boxes overlap, then the clearance refers to the number of overlapping pixels with negative symbol.

The comparison results are shown in Figures 8–9. Before character rectification, the text lines are inclined with various degrees because of the existence of perspective deformation in the character area. The character clearances are typically small, and the fluctuation range is big. When the perspective deformation is serious, overlaps are observed between the bounding boxes, thereby resulting in a large amount of negative clearance values. After deformation correction, the inclination angle of the text line is limited within 1°, and the character separations are also converged within the normal range. This finding substantiates that character images can be well deskewed and accurately cropped with our rectification algorithms.

C. EVALUATION ON RECOGNITION METHOD

To validate the influence of character rectification on recognition, we further study the accuracy of these three classifiers on the distorted dataset and the undistorted dataset. To make full use of datasets, we use five-fold cross-validation to test the model performance and calculate the average of the five results. For CRNN and ASTER, the image can be directly inputted for testing. For our CNN, we need to use the aforementioned MSER to segment the image before testing the model (Figure 6f). The results are shown in Table 1. These classifiers tested with undistorted images show remarkable performance, mainly because the deformation of characters and the overlap of bounding boxes before rectification bring greater pressure on the classifiers and drag down their accuracy rate. This finding also demonstrates that the performance of the classification algorithm can be greatly improved by our proposed rectification method. Notably, our CNN can reduce the error rate by 2 orders of magnitude with the undistorted images. In addition, we calculate the time cost of the three methods and the comparative result is shown in Table 2. As can be seen, our method is time saving and only based on CPU, while the deep networks of CRNN and ASTER require more time to complete the complicated computation.

Combined with the MSER method, our CNN outperforms the others on both datasets. Although the CRNN considers the context of the characters, the high randomness of wafer

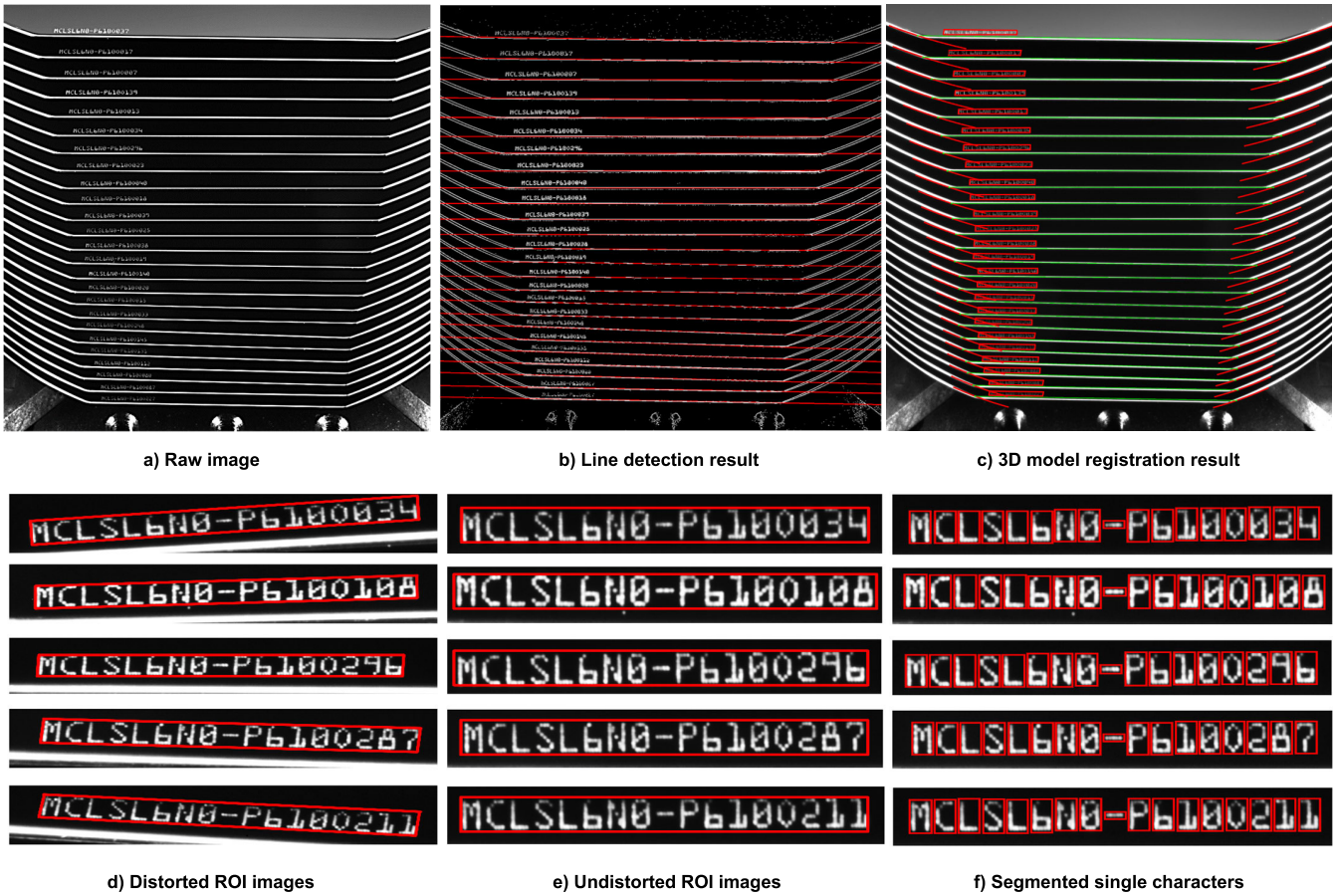


FIGURE 6. Result of each processing step. a) shows the raw image captured by the image acquisition scheme. b) illustrates the detected lines with the preprocessing method. c) demonstrates the 3D model registration result, where the red arc represents the searching angle for each wafer and green line represents the matched edge line of each wafer. d) shows the wafer-ID images cropped with the 3D model and detected angles. e) illustrates the undistorted ROI image after perspective transformation. f) demonstrates the segmented single characters of each wafer-ID.

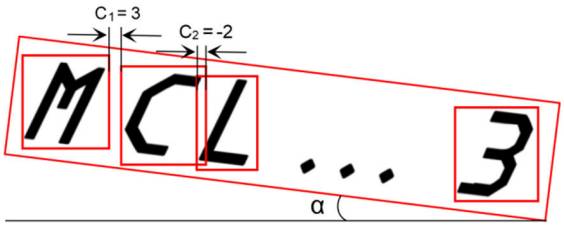


FIGURE 7. Bounding boxes of single character and text line.

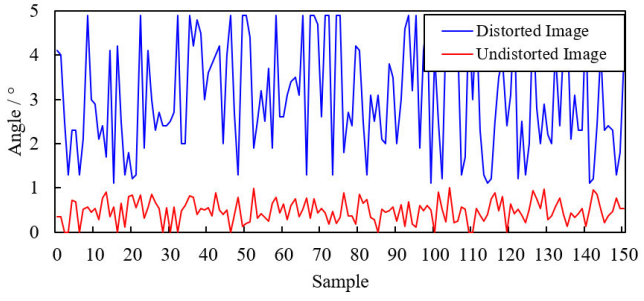


FIGURE 8. Tilt angle of text line.

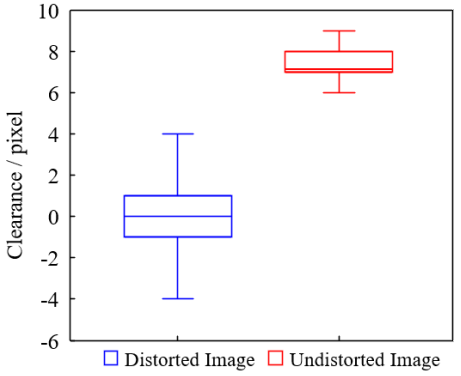


FIGURE 9. Box plot of character clearance before and after rectification.

IDs makes the improvement of accuracy difficult. The CRNN attempts to learn higher-level features for classification and adapts to various deformation through abundant training. However, due to the local change in illumination in the image, such as the glare and gloom, as well as various distortion of characters, learning enough features for classification with limited samples becomes difficult for the network. As for ASTER, which uses control points to adapt to deformation, it can slightly correct the image and show higher accuracy than CRNN. Nevertheless, the effect is very limited. In addition, word length affects the accuracy. As concluded in [14], a drop in accuracy is observed when the word length is higher than

TABLE 1. Accuracy comparation on two datasets.

Algorithm	Distorted Image-Dataset	Undistorted Image-Dataset
CRNN	64.74%	73.63%
ASTER	82.35%	88.24%
MSER+CNN	**86.94%**	99.98%

TABLE 2. Time-cost comparation on two datasets.

Algorithm	Time Cost (s)	Processor
CRNN	0.165	GPU
ASTER	0.574	GPU
MSER+CNN	**0.056+0.021**	CPU

11, because long words are inherently more difficult to predict correctly under the measure of whole-word accuracy. In our work, the word length of a wafer ID is 17, which is too long for the algorithms and requires more sufficient training datasets to achieve higher accuracy. As a result, our proposed 3D model registration-based undistortion method turns out to be so effective and necessary to allow the light-weight classifier to achieve high performance.

D. OVERALL EVALUATION ON SYSTEM

In the above experiments, we have verified all the functions of our proposed system, including image acquisition, preprocessing, segmentation, correction, and recognition algorithms, and constructed datasets to evaluate the algorithms. As shown in Figure 6a, the image acquisition scheme in our proposed system can obtain the clear image of batch wafers in a cassette at one time. The red line in Figure 6b represents the detected edge line of each wafer, which overlaps with the real edge line in the image. The red arcs in Figure 6c represent the searching angles of the 3D model registration algorithm. The rotation angle of each wafer can be searched by constantly projecting the end points of each wafer from the 3D model to the image to match the detected line. The searched edge lines are visualized in green in Figure 6c, thereby proving that the 3D model registration method can obtain the rotation angles of wafers accurately. Furthermore, according to the priori position of the ID characters on a wafer, the bounding boxes of wafer-IDs, which are represented by red quadrilaterals (Figure 6c), can be calculated using the searched angles. Then, perspective transformation is used to correct the wafer-ID images properly and the effectiveness of the correction algorithm is demonstrated (Figures 6d–e). Finally, as shown in Figure 6f, the single wafer-ID image is segmented using our proposed MSER method and inputted to the CNN for recognition results.

To further evaluate the overall performance of our proposed system, the time cost of each step in our system architecture is recorded for statistics and analyses. As shown in Table 3, we count the duration for preprocessing the entire image, undistorting 25 wafer-ID areas, and recognizing 25×17 single-character images separately. The total time consumption for recognizing a cassette of wafers with our method is within 4 s, which is less than one fifth of the recognition period of the equipment available on the market [8]–[10]. To sum up, our proposed system is competitive in both accuracy and efficiency.

TABLE 3. Time-cost of each part of our method.

Algorithm	Time Cost (s)
Preprocessing	0.448
Undistortion	1.471
Recognition	1.937
Total	3.856

VI. CONCLUSION

In this paper, we propose a batch wafer-ID recognition system and a 3D-model registration-based recognition algorithm. We design a specific image-acquisition system with linear-light sources and a high-resolution camera. Using this system, we focus on recognizing the IDs of a six-inch wafer. A series of processing steps for image preprocessing, undistortion, segmentation, and recognition are proposed. In addition, many pictures are collected to make datasets for model evaluation. The experimental results show that our method achieve high efficiency and high accuracy of batch wafer-ID recognition, thereby outperforming other state-of-the-art methods. In the future, we will further improve the efficiency and accuracy of our system through model optimization and parallel operation. Moreover, the system can be further implemented with more specifications of wafers, enriching the functions and applications.

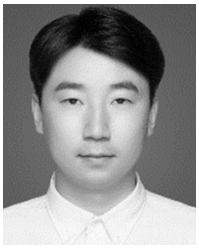
REFERENCES

[1] E. D. Williams, R. U. Ayres, and M. Heller, "The 1.7 kilograms microchip: Energy and material use in the production of semiconductor devices," *Environ. Sci. Technol.*, vol. 36, no. 24, pp. 5504–5510, 2002.
[2] M. Liukkonen and T.-N. Tsai, "Toward decentralized intelligence in manufacturing: Recent trends in automatic identification of things," *Int. J. Adv. Manuf. Technol.*, vol. 87, nos. 9–12, pp. 2509–2531, 2016.
[3] J. Zuo and E. Darici, "A robust table registration method for batch table OCR processing," in *Proc. 4th Int. Workshop Multilingual (OCR–MOCR)*, 2013, pp. 1–5.
[4] X. Fang, X. Fu, and X. Xu, "ID card identification system based on image recognition," in *Proc. 12th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2017, pp. 1488–1492.
[5] H. Li, P. Wang, M. You, and C. Shen, "Reading car license plates using deep neural networks," *Image Vis. Comput.*, vol. 72, pp. 14–23, Apr. 2018.
[6] I.-M. Ahn, "Development of a multi-template type image segmentation algorithm for the recognition of semiconductor wafer ID," *Trans. Korean Inst. Elect. Eng. P*, vol. 55, no. 4, pp. 167–175, 2006.
[7] W.-C. Hsu, T.-Y. Yu, and K.-L. Chen, "Robust wafer identification recognition based on asterisk-shape filter and high-low score comparison method," *Appl. Opt.*, vol. 48, no. 35, pp. 6606–6620, 2009.
[8] *Sorter OCR 150/200*. Accessed: Nov. 7, 2021. [Online]. Available: https://www.r2d-automation.com/product/sorter-150-200
[9] *Wafer ID Reader: WID-110*. Accessed: Nov. 7, 2021. [Online]. Available: https://www.ioss.de/en/code-reading-systems/fixmount-reader/wafer-identification-reader

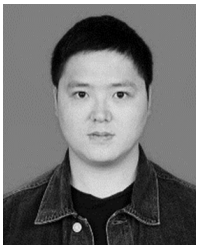
[10] *ID-Scope Automatic Wafer ID Reader*. Accessed: Nov. 7, 2021. [Online]. Available: https://www.glautomation.com/index.php?option=com_content&task=view&id=21&Itemid=35
[11] J. Friedman *et al.*, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2017.
[12] Z. Tian, W. Huang, T. He, P. He, and Y. Qiao, "Detecting text in natural image with connectionist text proposal network," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 56–72.
[13] Y. Liu and L. Jin, "Deep matching prior network: Toward tighter multi-oriented text detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1962–1969.
[14] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 11, pp. 2298–2304, Nov. 2017.
[15] P. He, W. Huang, Y. Qiao, C. C. Loy, and X. Tang, "Reading scene text in deep convolutional sequences," in *Proc. Conf. Artif. Intell.*, 2016, pp. 3501–3508.
[16] C. Bartz, H. Yang, and C. Meinel, "STN-OCR: A single neural network for text detection and text recognition," 2017, *arXiv:1707.08831*.
[17] B. Shi, M. Yang, X. Wang, P. Lyu, C. Yao, and X. Bai, "ASTER: An attentional scene text recognizer with flexible rectification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 9, pp. 2035–2048, Sep. 2019.
[18] L. Qiao, S. Tang, Z. Cheng, Y. Xu, Y. Niu, S. Pu, and F. Wu, "Text perceptron: Towards end-to-end arbitrary-shaped text spotting," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 7, 2020, pp. 11899–11907.
[19] J. Li, Z. Zhou, Z. Su, S. Huang, and L. Jin, "A new parallel detection-recognition approach for end-to-end scene text extraction," in *Proc. Int. Conf. Document Anal. Recognit. (ICDAR)*, Sep. 2019, pp. 1358–1365.
[20] J. Duan, Y. Xu, Z. Kuang, X. Yue, H. Sun, Y. Guan, and W. Zhang, "Geometry normalization networks for accurate scene text detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 9137–9146.
[21] Y. Liu, H. Chen, C. Shen, T. He, L. Jin, and L. Wang, "ABCNet: Real-time scene text spotting with adaptive Bezier-curve network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 9809–9818.
[22] *Specification for Alphanumeric Marking of Silicon Wafers*, Standard GB/T 34479, National Standard of the People's Republic of China, 2017.
[23] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.
[24] D. Forsyth and J. Ponce, "Image formation," in *Computer Vision: A Modern Approach*. New York, NY, USA: Prentice-Hall, 2011.
[25] M. Donoser and H. Bischof, "Efficient maximally stable extremal region (MSER) tracking," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2006, pp. 553–560.
[26] A. Neubeck and L. Van Gool, "Efficient non-maximum suppression," in *Proc. Int. Conf. Pattern Recognit.*, vol. 3. Aug. 2006, pp. 850–855.
[27] El-Sawy, E.-B. Hazem, and M. Loey, "CNN for handwritten Arabic digits recognition based on lenet-5," in *Proc. IEEE Int. Conf. Adv. Intell. Syst. Inf. (AISI)*, 2016, pp. 566–575.
[28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[29] M. D. Zeiler, "ADADELTA: An adaptive learning rate method," 2012, *arXiv:1212.5701*.

FANG CAO was born in 1995. She received the B.S. degree in measurement and control technology from the Nanjing Institute of Engineering, Nanjing, China, in 2018. She is currently pursuing the M.S. degree with the School of Physics and Microelectronics, Zhengzhou University, Zhengzhou, China. Her current research interests include machine vision and machine learning.

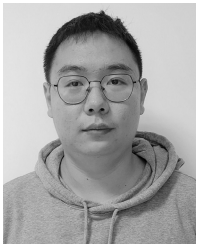
ZENGGUO TIAN was born in 1971. He received the M.S. degree in mechanical and electronic engineering from the Wuhan University of Science and Technology, Wuhan, China, in 2002. He is currently an Associate Professor with the School of Physics and Microelectronics, Zhengzhou University, Zhengzhou, China. His research interests include industrial inspection, semiconductor material, and intelligent manufacturing.



BAOZHU JIANG was born in 1996. He received the B.S. degree in measurement and control technology and instruments from Zhengzhou University, Zhengzhou, China, in 2018, where he is currently pursuing thes M.S. degree in instrument engineering with the School of Physics and Microelectronics. His research interests include computational imaging and machine vision.



HONGSHUAI ZHANG was born in 1991. He received the M.S. degree in mechanical engineering from the School of Mechanical and Power Engineering, Zhengzhou University, Zhengzhou, China, in 2016, where he is currently an Assistant Experimentalist with the School of Mechanical and Power Engineering. His research interests include intelligent manufacturing and industrial internet.



HENG CHEN was born in 1996. He received the B.S. degree in electrical engineering and automation electrical engineering from Zhengzhou University, Zhengzhou, China, in 2020, where he is currently pursuing the M.S. degree in instrument engineering with the School of Physics and Microelectronics. His research interests include image processing and pattern recognition.

XUGUANG ZHU was born in 1988. He received the M.S. degree in material processing engineering from the School of Materials Science and Engineering, Henan University of Science and Technology, Luoyang, China, in 2015. He is currently a Material Engineer at Luoyang Dejing Intelligent Technology Company Ltd. His research interest includes semiconductor thermoelectric material.