Multimedia Tools and Applications (2021) 80:32409–32421
https://doi.org/10.1007/s11042-021-10981-2

Zernike-CNNs for image preprocessing and classification in printed register detection

Sheng Wang[1] · Lin-Tao Lv[1] · Hong-Cai Yang[1] · Di Lu[2]

Received: 15 September 2020 / Revised: 29 December 2020 / Accepted: 30 April 2021 /
Published online: 31 July 2021

Abstract
In the register detection of printing field, a new approach based on Zernike-CNNs is proposed. The edge feature of image is extracted by Zernike moments (ZMs), and a recursive algorithm of ZMs called Kintner method is derived. An improved convolutional neural networks (CNNs) are investigated to improve the accuracy of classification. Based on the classic convolutional neural network (CNN), the improved CNNs adopt parallel CNN to enhance local features, and adopt auxiliary classification part to modify classification layer weights. A printed image is trained with 7×400 samples and tested with 7×100 samples, and then the method in this paper is compared with other methods. In image processing, Zernike is compared with Sobel method, Laplacian of Gaussian (LoG) method, Smallest Univalue Segment Assimilating Nucleus (SUSAN) method, Finite Impusle Response (FIR) method, Multi-scale Morphological Gradient (MMG) method. In image classification, improved CNNs are compared with classical CNN. The experimental results show that Zernike-CNNs have the best performance, the mean square error (MSE) of the training samples reaches 0.0143, and the detection accuracy of training samples and test samples reached 91.43% and 94.85% respectively. The experiments reveal that Zernike-CNNs are a feasible approach for register detection.

Keywords Printed defective detection · Register measurement · Edge detection · Zernike moments (ZM/ZMs) · Convolutional neural networks (CNN/CNNs) · Pattern recognition and classification

1 Introduction

Register is defined as the accuracy of the relative positions between any two color images in polychrome printing. In the transfer process, if the image information deviates from the

✉ Sheng Wang
ws5000@sina.com

[1] Xijing University, Xi'an 710123 Shaanxi, China

[2] Xi'an University of Architecture and Technology, Xi'an 710055 Shaanxi, China

predetermined position, it is called “register error”, “register difference”, “mis-register” or” out of register”, which is a kind of common printing trouble. Figure 1 shows examples of register errors in several printed materials:

As can be seen from Fig. 1, the register errors in printed matters are usually very smooth and the features are not obvious. The traditional manual inspection mainly inspects the register lines, which are cross lines. When the register marks deviation is too large, it is considered that there are register errors, then checked candidate images to find out the defective images. The cross lines for register are shown in Fig. 2.

Manual inspection for register is labor intensive and unreliable [27]. Automatic detection instead of manual is the development trend. At present, there are relevant facilities on the market for printing flaw detection and even register measurement, such as Nota SaveCheck in Switzerland, ErgoTronic series and QualiTronic from Koenig-Bauer(KBA) in Germany. Perhaps due to the reasons of commercial confidentiality, the public algorithms of quality inspection equipment from each company have not been found. Although it is not clear about the details of quality inspection equipment on the market, register detection methods can be generally inferred to adopt two directions. One is to detect the coincidence of cross lines, which is a traditional method. The other method is to detect images directly based on computer vision, which has a high degree of automation and precision.

The main problems of existing register control, measurement and defect detection can be listed as follows:

(1). Cross lines are mainly used for manual inspection, with high labor intensity and unstable work quality. The printed matters usually be printed several same small formats on a press sheet. Although the press sheet is not matched properly, due to the scalability of the paper, the deviation of individual small images may be allowed in a certain level of printing quality. At the other extreme, there will be cross is accurate, but the small images are not accurate. At these time, it is necessary to manually inspect the small sheet products which achieve the standard or not, which will increase the labor intensity and reduce the detection efficiency.

(2). Computer vision, the influence of noise is great and feature extraction is difficult. This problem belongs to image preprocessing. The detection effect of register measurement depends on the position features of image contours. However, due to the light, camera (sensors) accuracy and shooting angle, image digitization and other reasons, it is often

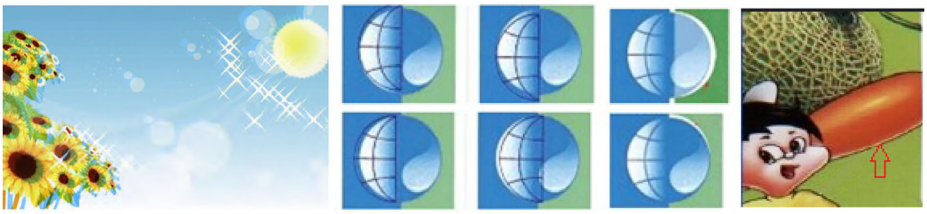


Fig. 1 Register Errors in Printed Matters: (**a** adapted from Fig. 3 of [37]) The whole picture is a double image due to register error, and the printing defect is the most obvious in the red box; (**b** adapted from Fig. 2 of https://wenku.baidu.com/view/d56bdfbd240c844769eaee78.html) Several register errors on stamp patterns, The longitude and latitude lines in the left and middle pictures deviate from the earth, In the right figures, the white background is leaked due to register deviation;(**c** adapted from Fig. 1 of http://ebook.keyin.cn/magazine/pt-bzzh/201208/27-972928.shtml) Arrow indication: register error between color and edge

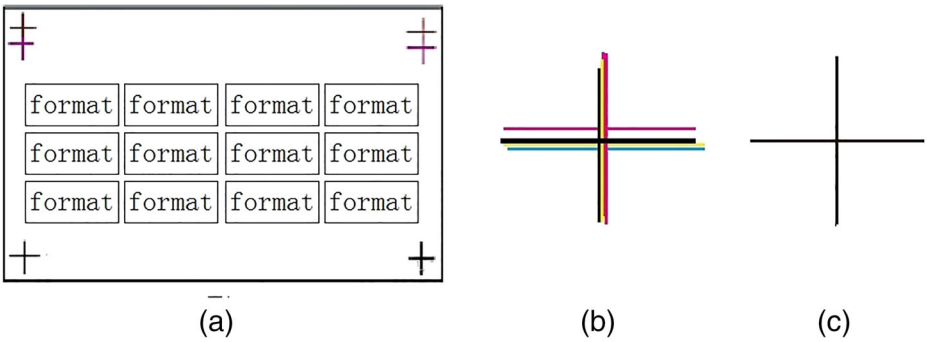


Fig. 2 Cross Lines for Register: (**a**) The layout of cross lines is located at the four corners of the press sheet; (**b**) Cross lines are composed of cyan, magenta, yellow and black, if there are errors, cross lines do not coincide (**c**). When register is accurate, cross lines overlap and appear black

difficult to obtain the ideal position information. Therefore, filtering noise, location feature extraction and amplification of printed image are the key points for register measurement in the process of image preprocessing.

(3). High false error rate from computer vision. This problem belongs to image classification. Existing high-precision algorithms for automatic register measurement are complex. Nevertheless, there are still missing and false detection. Missing detection is understood to the defective image is judged as good images, while false detection is understood to the good image is judged as defective image. In some special printing industries, such as maps, certificates, banknotes and security printing, quality security must be assured. The circulation of defective printed matters in the market will cause serious political, economic and social impact. Therefore, in order to avoid missing detection of defective printed products, the strict thresholds are often set by special printing enterprises in machine detection. Although this kind of practice can avoid the defective printed products being put into the market to a certain extent, it also makes many qualified printed products be intercepted as defective, resulting in the enhance of false error rate, increasing the waste and cost of enterprises. According to the statistics from a certain type machines for quality inspection of printed products, the false error rates in classification fluctuate between 6.8% and 21.4% for every 100,000 sheets, which is still considerable.

(4). Cost of computer vision is high. Computer vision needs performance hardware, which leads to high investment; Computer vision still needs a lot of time for sample learning, during which the detection equipment can not be applied to the productive process, resulting in a blank window period. These factors lead to higher costs.

Based on the above analysis and description, this paper investigates the register control and measurement from the perspective view of image recognition. The main contributions of this paper include:

(1). A new approach, denoted as Zernike-CNNs method, is proposed. It is a new research region to apply computer vision for printing register detection. Computer vision is more advanced and reliable than cross lines. This method can not only distinguish the standard products and defective products, but also subdivide the types of defective products, which can satisfy the needs of defective analysis. The steps are: Color images are

converted to CMYK mode; The Y-gray image of printed matter is preprocessed by Zernike moments (ZMs), and then the convolution neural networks (CNNs) are used for image recognition.

(2). In order to overcome the disadvantage of complex factorial calculation for ZMs radial polynomials, a fast recurrence algorithm, called Kintner method for Zernike moment radial polynomials, is derived in this paper. In order to further improve the classification accuracy, an improved CNNs is proposed. On the basis of CNN, paralleled CNN for local feature extraction and auxiliary classification for classification layer are proposed.

(3). Zernike method (ZM) is compared with Sobel method, Laplacian of Gaussian (LoG) method, Small Univalue Segment Assimilating Nucleus (SUSAN) method, Finite Impulse Response (FIR) method, Multi-scale Morphological Gradient (MMG) and other preprocessing methods, The results show that ZMs belongs to integral operation, which are good behaviors in noisy conditions; it can locate at sub-pixel level with high accuracy; ZMs can be good candidates for rotational and scaling images, which is very beneficial to the detection of images collected with various deformations by the printing press at high-speed.

(4). Zernike-CNNs method mainly detects the register defects of printed matter, as well as has certain detection ability for other types of printing defects. This approach is more adaptable to printing defective detection.

The rest of this paper is organized as follows: Section 2 introduces the academic research on register control, measurement and computer vision; In Section 3, the background information of the research methods in this paper is described, a fast Zernike algorithm is derived, and proposes an improved CNNs. Section 4 presents the technique route for register measurement. The experimental results are analyzed and discussed in Section 5. The conclusions and some future research activities are drawn in Section 6.

2 Related works

The academia has carried out a lot of research to improve the effect of automatic register measurement and control. Yoshida [39, 40] established the mathematical model of the interaction between upstream and downstream register control, and demonstrated the validity of the model through experiments. Lee [14, 22–24] illustrated the influence of thermal, register marks, edge-detection and algorithm on register control and measurement. On this basis, the optimization method on the register measurement error is studied. Kang [16, 17] proposed a prediction models for cross directional (CD) and machine directional (MD) register in multilayer printing, and verified the correctness of the models through experiments. For improving the register performance, Kim [19] proposed the register measurement system and related control method. Sugimoto [33] developed a register controller on the gravure press in response to the construction of the printing press and printing conditions such as the pass length between the printing units, the compensator roller displacement velocity, the printing speed, so as to achieve the purpose of register control. These studies focused on the prediction and prevention of register error, but do not involve the detection of register error. Therefore, these methods can not timely and completely correct the causes of defects in the printing process, which is easy to form tendentious defective products.

In the field of printing, the relevant academic papers and achievements on quality inspection are relatively rare. In fact, printing quality inspection can be regarded as the category of computer vision. Computer vision is widely used in authentication [9, 31], classification [38], inspection and detection [34], measurement and positioning [30], and other fields. There are many researches related to computer vision, the researches on computer vision have important reference significance for the detection on printing quality defects. The researches of computer vision mainly focus on:

(1) Preprocessing of sample data set, which aims to optimize the calculation amount of data, reduce noise interference and enhance data features.

Dixit [9] proposed a forgery detection method for digital images based on stationary wavelet transform (SWT) with singular value decomposition (SVD), However, this method did not solve the problems of forgery detection after image rotation, rescale and reflection. Yasaka [38] studied the analysis accuracy of different CT images with Laplacian of Gaussian (LoG) spatial band-pass filter. However, when LoG enhances image features, it is easy to bring noise to image edge profile. Verma [34] proposed smallest univalue segment assimilating nucleus (SUSAN) principle and bacterial foraging algorithm (BFA) for edge detection. The performance of SUSAN operator is not affected by template size, and the selection of parameters is very simple. Compared with other edge detection algorithms, this algorithm was proved to be more effective by t-test. But SUSAN is mainly used for corner detection and has poor anti-noise ability, so it is not suitable for edge smoothed images. Müller [30] used the Sobel method to process deep optical images of satellite galaxy. But this method is still differential filtering, which is easy to amplify noise and produce false contour. Lupek [28] used finite impulse response (FIR) filter to remove spectral noise and background broadband distortion in Raman spectra. The experimental results show that FIR filter is an effective method to process the whole Raman spectra. But this method has a large amount of calculation. Diop [8] proposed to use the image edge functions as the local weights of inhomogeneous Hamiltonians to obtain multi-scale morphological operators. The method has good performance in image processing, but the algorithm can not completely avoid the situation of over segmentation. Firuzi [12] mentioned Histogram of Oriented Gradient(HOG), Kumar [21] mentioned Scale-Invariant Feature Transform(SIFT), These algorithms extract image features using artificially designed image descriptors, The algorithms can only be used for specific objects and are not universal.

Moments are a common method for image processing. The commonly used moments are Hu moments and Zernike moments (ZMs). Fernando [11] thought Hu moment is the best method of image processing. Li [26] proposed sub-pixel edge location based on Zernike with the error function edge model. The method has good robustness and location accuracy, but the improved algorithm increases the amount of computation. In addition, the algorithm must change the edge model as well as different objects. Compared with Hu moments, ZMs have the following advantages: it is easy to construct arbitrary high-order moments; it is not sensitive to noise.

(2) Algorithm researches on data feature recognition and classification. The commonly used methods are classical methods and deep-learning-based methods [7].

Support vector machine (SVM) is a common classifier. Vidi [35] applied SVM to predict breast tumors benign versus malignant, classify breast cancer subtypes, and achieved good

results. In the fingerprint identification technology, Kaur [18] combined ZMs with SVM, used ZMs to extract the features of fingerprint image. Then, a weighted-SVM is used to train and test the evaluated features. Experimental results show that the performance of this method is significantly improved compared with the existing technology. SVM is more suitable for binary classification, and the applicability for multi classification remains to be observed. Parameters selection is also the main constraint factor for the effect. In addition, there are other classification algorithms. CHEN [4] mentioned k-nearest neighbor(KNN) algorithm and naïve Bayes classifier(NBC) in behavior classification, and achieved better results. However, the calculation of KNN leads to long processing time, each sample to be classified has to calculate its distance from all known samples in order to get its k-nearest neighbor. NBC needs the assumption that the feature conditions are independent, otherwise the classification accuracy will be reduced.

Deep learning has excellent performance in image classification. Lu Leng [25] proposed a threshold segmentation scheme to preprocess the images of biological samples, and used a light-weight shallow CNN to classify the images. This method improves the computational efficiency of image detection and classification. At present, neural network represented by deep learning has become the main method of image classification.

In summary, register measurement and control belongs to computer vision, which needs image preprocessing and classification. In the preprocessing stage, because register detection is to extract the subtle edges from image, wavelet and other texture feature extraction algorithms are not applicable to register measurement; As the image edges related to register are generally smooth, so HOG, SIFT and other local feature extraction algorithms are not applicable to this object; because the sheets are running at a high speed in the printing press, the collected images are affected by light, angle, paper tension and other factors, there are many noise, differential type filters will suffer from well performing under the noisy conditions. ZMs are an integral operation with invariance and strong noise against ability, and ZMs can extract edges at sub-pixel level, so ZMs are adopted as a preprocessing method for register detection. In this paper, compared the integral operation represented by Zernike with the differential filter represented by Sobel, second order differential represented by LoG, the circular template represented by SUSAN, the wavelet is represented by FIR, nonlinear filter represent by MMG to verify their performance. In the classification stage, because the printed matter should not only distinguish good products from defective products, but also distinguish the types of defective products, so it belongs to multi classification. The binary classification of SVM has a large amount of calculation, and the classification effect is not good. Therefore, CNN, represented by deep learning, is adopted to classify the printed in this paper.

3 Approach

3.1 CMYK mode

The four colors offset pressing uses cyan (C), magenta (M), yellow (Y) and black (K) as the basic colors, and realizes color printing according to the subtractive process. The printing image obtained by the detection unit is generally in RGB space, so the image data need to be converted into the CMYK space under color separations [10, 15]. Among them, C-M-Y gray values are complementary to the gray value of red (R), green (g), blue (b). Due to the gray values range of RGB [0,255], the extraction process of C-M-Y gray value is given by

$$\begin{cases} C_{ij} = 255 - R_{ij} \\ M_{ij} = 255 - G_{ij} \\ Y_{ij} = 255 - B_{ij} \end{cases} \tag{1}$$

where C_{IJ}, M_{ij}, Y_{ij} and R_{ij}, G_{ij}, B_{ij} are the gray values of pixels under different color separations. Then the C-M-Y components are extracted to form black (k) components, and the C-M-Y components are modified. The correction process is given by

$$\begin{aligned} &K_{ij} = \min\left(C_{ij}, M_{ij}, Y_{ij}\right) \\ &if \quad K_{ij}! = C_{\mathrm{K}} \\ &\begin{cases} C_{ij} = \alpha \times \left(C_{ij} - K_{ij}\right) / \left(C_{\mathrm{K}} - K_{ij}\right) \\ M_{ij} = \alpha \times \left(M_{ij} - K_{ij}\right) / \left(C_{\mathrm{K}} - K_{ij}\right) \ else C_{ij} = M_{ij} = Y_{ij} = 0 \\ Y_{ij} == \alpha \times \left(Y_{ij} - K_{ij}\right) / \left(C_{\mathrm{K}} - K_{ij}\right) \end{cases} \end{aligned} \tag{2}$$

where C_K and α are empirical values. In this paper, C_K is 100 and α is 1. According to the above Eq.(1 ~ 2), the gray value extraction in CMYK mode is completed.

3.2 Zernike moments

Edge detection of candidate image can effectively enhance image features, recognize patterns, and reduce the amount of computation. Theoretical research shows that ZMs are an integral form, which have good noise tolerant ability, and can construct any order of orthogonal moments in the unit circle to convolute the image [2]; The ZMs have translation-, rotation- and scale-invariance [3], so the ZMs are a common edge detection algorithm [13]. The Zernike moment (ZM) is defined as

$$Z_{nm} = \frac{n+n}{\pi} \iint_{x^2+y^2 \le} f(x,y) * V_{nm}^*(x,y) dxdy \quad = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho,\theta) * V_{nm}^*(x,y) \rho d\rho d\theta \tag{3}$$

where Z_{nm} is the m-fold repetition of n-th moment for Zernike, and N-$|m|$ is even, $n \ge |m|$. *f(x, y)* is the pixel value of the image at *(x, y)* pixel point. $V_{nm}{}^*$ is the conjugate of Zernike polynomials. Z_{nm} is obtained by convolution of the pixel value *f(x, y)* of pixel *(x, y)* with V_{nm}*. For an image with $N \times N$ pixels, Zernike polynomials V_{nm} is defined as

$$\begin{aligned} &V_{nm}(x,y) = R_{nm}(\rho) \exp(jm\theta) \\ &\begin{cases} R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} \dfrac{(-1)^s (n-s)! \rho^{n-2s}}{s! \left(\dfrac{n+|m|}{2} - s\right)! \left(\dfrac{n-|m|}{2} - s\right)!} \\ \rho = \dfrac{\sqrt{(2x-N+1)^2 + (N-1-2y)^2}}{N} \\ \theta = \tan^{-1}\left(\dfrac{N-1-2y}{2x-N+1}\right) \end{cases} \end{aligned} \tag{4}$$

where $R_{nm}(\rho)$ is the Zernike radial polynomials, ρ is the distance from the reference point to the pixel, and θ is the counter clockwise angle of the pixel relative to the positive direction of the x-axis. The relationship between the ZMs before and after rotation be shown that

$$Z'_{nm} = Z_{nm} e^{-jm\varphi} \tag{5}$$

where Z_{nm}' is the ZM after an image rotated angle φ. The Eq. (5) proves the rotational invariance of the ZMs. When the image pixel is rotated around a reference point, the module value of the pixel relative to the reference point remains unchanged except for the change of the phase angle. Some important information of the image can be determined by ZMs before and after rotation, and the information expression is shown as

$$\begin{cases} \varphi = \tan^{-1} \dfrac{\mathrm{Im}|Z_{11}|}{\mathrm{Re}|Z_{11}|} \\ l = \dfrac{Z_{20}}{Z'_{11}} \\ k = \dfrac{3Z_{11}\exp(-j\varphi)}{2\left(1-l^2\right)^{3/2}} \\ h = \dfrac{Z_{00} - \dfrac{k\pi}{2} + k\sin^{-1}(l) + kl\sqrt{1-l^2}}{\pi} \end{cases} \tag{6}$$

Where φ is the phase angle of the pixel rotation, l is the distance between the reference point and the image contour, k is the gray step height on both sides of the contours, and h is the background gray level. If the pixel satisfies the requirements of $K \geq k_t \cap l \leq l_t$, where K_t and l_t are the threshold values, then it can be judged that the pixel *(x, y)* is the contour of the image. The selection of threshold has a great impact on the image edge location. l_t is generally set to be less than 0.5, and K_t needs to be set experimentally or empirically.

According to Eq.(4), the factorial terms of radial polynomial $R_{nm}(\rho)$ causes the complexity of ZMs, which hinders the application of ZMs in real time [6]. In order to reduce the burden of these factorial terms, Wee [36] and Mukundan [29] have, respectively, introduced several fast computations of the ZMs, for example: Kintner and Prata methods, and so on. However, the derivation process of the fast computation is not given in these literatures. Chen [5] proposed the application of Riemann sum to approximate Zernike radial function.

In the previous discussion, a fast algorithm of Zernike radial polynomials, which is called Kintner method, is very famous, but it has not seen to deduce this method by any article. In this paper, Kintner method is derived.

Zernike radial polynomials can be regarded as a special case of Jacobi polynomials. There is a corresponding relationship between Zernike radial polynomials and Jacobi polynomials as follows:

$$R_{nm}(\rho) = \rho^m P^{(0,m)}_{\frac{1}{2}(n-m)}\left(2\rho^2-1\right) \tag{7}$$

Where $P^{(9,m)}_{\frac{1}{2}(n-m)}\left(2\rho^2-1\right)$ are Jacobi polynomials. Jacobi polynomials have a recurrence relations as shown below

$$\begin{aligned} &2n^*\left(\alpha+\beta+n^*\right)\left(\alpha+\beta+2n^*-2\right)P^{(\alpha,\beta)}_{n^*}(x) = \\ &\left(\alpha+\beta+2n^*-1\right)\left[\alpha^2-\beta^2+x\left(\alpha+\beta+2n^*\right)\left(\alpha+\beta+2n^*-2\right)\right]P^{(\alpha,\beta)}_{n^*-1}(x) \\ &-2\left(\alpha+n^*-1\right)\left(\beta+n^*-1\right)\left(\alpha+\beta+2n^*\right)P^{(\alpha,\beta)}_{n^*-2}(x) \end{aligned} \tag{8}$$

Let $n^* = \frac{1}{2}(n-m), \alpha = 0, \beta = m$, Using the Eq. (7 ~ 8) have the following relation

$$\begin{aligned}&\frac{1}{2}(n-m)(n+m)(n-2)P^{(0,m)}_{\frac{1}{2}(n-m)}(x) = \\ &(n-1)\left[-m^2+\left(2\rho^2-1\right)n(n-2)\right]P^{(0,m)}_{\frac{1}{2}(n-m)-1}(x)-n(n-m-2)\left(\frac{1}{2}n+\frac{1}{2}m-1\right)P^{(0,m)}_{\frac{1}{2}(n-m)-2}(x)\end{aligned} \tag{9}$$

In Eq. (9), let $x = 2\rho^2 - 1$, Jacobi polynomials can be transformed into Eq. (10) from Eq. (7).

$$\begin{cases} P^{(0,m)}_{\frac{1}{2}(n-m)}(x) = P^{(0,m)}_{\frac{1}{2}(n-m)}\left(2\rho^2-1\right) = \rho^{-m}R_{n,m}(\rho) \\ P^{(0,m)}_{\frac{1}{2}(n-m)-1}(x) = P^{(0,m)}_{\frac{1}{2}(n-2-m)-1}\left(2\rho^2-1\right) = \rho^{-m}R_{n-2,m}(\rho) \\ P^{(0,m)}_{\frac{1}{2}(n-m)-2}(x) = P^{(0,m)}_{\frac{1}{2}(n-4-m)}\left(2\rho^2-1\right) = \rho^{-m}R_{n-4,m}(\rho) \end{cases} \tag{10}$$

The fast recurrence relation for ZM polynomials can be obtained from Eq. (9 ~ 10). See Eq. (11) for fast recurrence relation:

$$\begin{aligned}&\frac{1}{2}(n-m)(n+m)(n-2)R_{n,m}(\rho) = \\ &(n-1)\left[-m^2+\left(2\rho^2-1\right)n(n-2)\right]R_{n-2,m}(\rho)-n(n-m-2)\left(\frac{1}{2}n+\frac{1}{2}m-1\right)R_{n-4,m}(\rho)\end{aligned} \tag{11}$$

However this relation as shown in Eq. (11) is not applicable for cases *(n = m)* and *(n-m = 2)*. For these cases, the following two relations are used:

$$\begin{aligned}&R_{n,n}(\rho) = \rho^n \\ &R_{n,n-2}(\rho)-nR_{n,n}(\rho)-(n-1)R_{n-2,n-2}(\rho)\end{aligned} \tag{12}$$

The direct method of Eq. (4) can be replaced by the fast computation method in Eq. (11 ~ 12) for the Zernike radial polynomials. Eq.(11 ~ 12) are the Kintner method.

When the Zernike radial polynomials are from 0 to the specified orders, and under the specified order, the efficiency comparison between the fast computation method and the direct method is shown in Table 1.

In Table 1, Kintner method takes shorter time than the direct method in Eq. (4). The higher the orders, the more efficient the fast computation method is.

3.3 Convolution neural network

Convolutional Neural Network (CNN) is an efficient deep learning algorithm. The network can extract the object features by local data. The network has certain translation-, rotation- and scale-invariance, and has been used to features recognition [1] and classification function [20]. The network can not only detect register measurement of printed matter, but also detect other defects, so it has a wide range of application prospect in the detection of printing defects.

This paper is based on Le net-5 CNN, and its architecture is shown in Fig. 3:

The components of a CNN are composed of several convolutional layers *(C1, C2…Cn)*, several pooling layers *(S1, S2…Sn)*, fully connected layer and classification layer. In general,

Table 1 Comparison between fast computation method and direct method

Methods orders Time(s)	direct method	Kintner method	Methods order Time(s)	direct method	Kintner method
n=0~4	0.366	0.144	n=4	0.135	0.065
n=0~6	0.738	0.391	n=6	0.220	0.098
n=0~8	1.414	1.046	n=8	0.471	0.188
n=0~10	2.581	2.613	n=10	0.706	0.214
n=0~20	18.260	7.129	n=20	2.601	0.616
n=0~30	58.675	13.600	n=30	5.334	1.214
n=0~40	128.010	29.759	n=40	9.012	2.259
n=0~50	232.960	57.241	n=50	13.702	3.381
n=0~60	400.469	99.969	n=60	19.183	5.096
n=0~70	696.566	162.220	n=70	26.740	7.112

the convolutional layer and pooling layer appear alternately in pairs. For the basic idea and operation of CNNs are given by Ref. [32].

One of the challenging tasks from CNN is that, a particular network cannot be applied to continuously increasing detection environments. The parameters and structure of the network must be improved according to the characteristics of the research object, or combine different algorithms to improve the work efficiency. This paper focuses on improving the classification accuracy of the network, an improved neural network is proposed. Its structure is shown in Fig. 4.

The improved CNNs include three parts: original part (OP), parallel part (PP) and auxiliary classification part (ACP).

The original part (OP) is classic CNN. For shorten the training time, the CNN in this paper has two convolution layers, but one pooling layer is reduced. Its architecture is: convolutional layer *C1*, pooling layer *S1*, convolutional layer *C2*, fully connected layer and classification layer. In the classic CNN, Softmax is used as classification function, and Mean square error (MSE) between network output y and expected output $\widetilde{y}$ as a loss function. MSE can be defined as:

$$E = d_1\left(y_m, \widetilde{y}_m\right) = \frac{1}{2m}\sum_{1}^{m}\left(y_m - \widetilde{y}_m\right)^2 \tag{13}$$

Where, m is the number of categories; the loss function of classification is calculated by gradient descent method to update the weight. But in the improved CNNs, the loss function from OP-CNN adopts a new model different from Eq. (13).

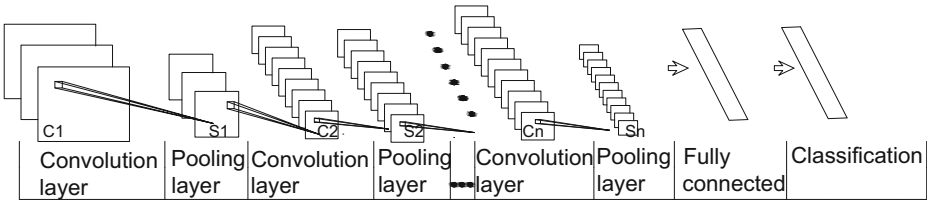


Fig. 3 Architecture of convolution neural network

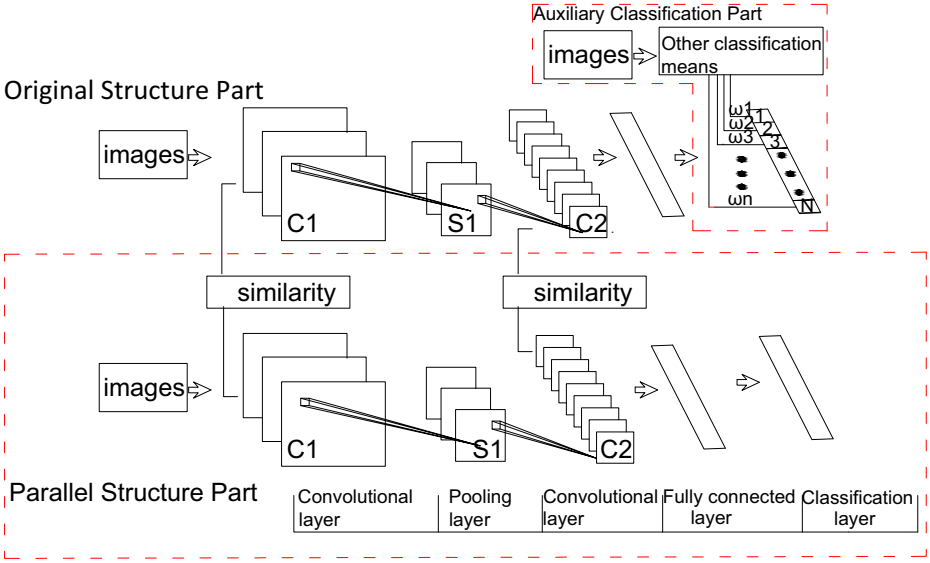


Fig. 4 Architecture of improved CNNs

Based on the original structure, the improved CNN adds two parts: parallel structure part and auxiliary classification part.

The parallel part (PP) has the same structure as the OP, Different from OP, PP's input image is local image, while OP's input image is global image. The relationship between OP and PP is established through the similarity function of output features from convolution layers. The similarity function and the output loss function constitute the training function for OP:

$$\begin{aligned} E_{op} &= \alpha \times d_1(y_m, \widetilde{y}_m) + \beta \sum_1^n d_2\left(C_{op,k}, C_{pp,k}\right) \qquad k = 1, 2 \ldots N \\ &= \frac{\alpha}{2m} \sum_1^m (y_m - \widetilde{y}_m)^2 + \frac{\beta}{n} \sum_1^n \sqrt{\sum_{pq} \left(z_{i,j} - \widetilde{z}_{ij}\right)^2_k} \\ &= E_1 + E_2 \end{aligned} \tag{14}$$

Where, the first term on the right-hand side of the representation is the output classification loss function; The second term on the right-hand side of the representation is the convolution outputs similarity function, $C_{op,k}$, $C_{pp,k}$ denote the outputs of OP and PP in the k-th *(k∈[1,N])* layer from convolution layer. Z_{ij}, $\widetilde{Z}_{ij}$ represents the values of $C_{op,k}$ and $C_{pp,k}$ in row i and column j respectively. α, β are the weights of two items respectively. The output loss function and the similarity function are used as the loss function of OP, and the weights of OP network are updated by gradient descent method. It is a well-known process that the updating formula of the weights and thresholds from the Eq. (13), as well as the first term on the right-hand side from the Eq. (14), can be obtained by the chain derivation rule. In the second term on the right-hand side from the Eq. (14), the chain derivation rule can also be used to update the weights and thresholds. The weights and thresholds update can be written as follows:

$$\frac{\partial E_{op}}{\partial \omega^k} = rot180\left(C_{op,k-1}\right) * \left(\frac{\partial E_1}{\partial C_{op,k}} + \frac{\partial E_{2,1}}{\partial C_{op,k}} + \ldots + \frac{\partial E_{2,n}}{\partial C_{op,k}}\right) \quad where \quad \frac{\partial E_{2,n}}{\partial C_{op,k}}$$
$$\frac{\partial E_{op}}{\partial b^k} = \sum_{p,q}\left(\frac{\partial E_1}{\partial C_{op,k}} + \frac{\partial E_{2,i}}{\partial C_{op,k}} + \ldots + \frac{\partial E_{2,n}}{\partial C_{op,k}}\right)$$
$$= 2\left(C_{op,k} - C_{pp,k}\right) \tag{15}$$

This structure will make CNNs pay more attention to the region of interest (ROI) of images, so as to improve the ability of image recognition and classification.

Aiming at improving the classification ability of CNNs, the classification layer weights are proposed as the auxiliary classification part. In some specific image recognition and classification, there are other means which have stronger ability than CNNs. These classification methods can be applied to CNNs as an auxiliary. This paper proposes that, the auxiliary classification algorithm classifies the candidate images, and the results are loaded into the classification layer of CNNs as weights to strengthen the correct classifications of CNNs and weaken the wrong classifications of CNNs.

In the next section, the detailed use of improved CNN in register detection will be described.

4 Realization of register measurement

4.1 Inspection system of register measurement

Machine for the automatic inspection and classification in this paper is shown in Fig. 5.

The system consists of four important units: LED light source, which provides continuous and stable illumination for the inspection system restraining noisy, reduces the deviation of detection result caused by the fluctuation of surrounding illumination; CCD sensor, collects printed matter images and inputs them to high-speed industrial computer; high-speed industrial computer, processes and classifies the collected printed matter, and coordinates the components of the detection system operation; PLC, complete the action control of detection system and data exchange with industrial computer.

4.2 Experimental study

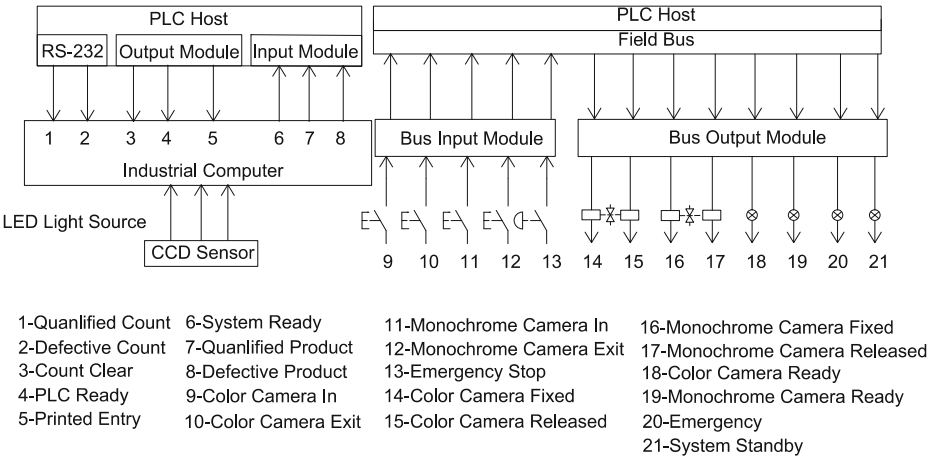
In this paper, the images to be checked for the inspection machine are taken as shown Fig. 6.

Figure 6(a)~(g) are divided into seven categories. In this paper, 500 sheets of each type are collected, and the total number of samples is 3500 as the raw data.

According to section 1.1, color separations of the images to be inspected are carried out in CMYK mode, and the gray scales are extracted as shown in Fig. 7.

Considering that there are only black and white pixels in Y gray image, in this paper, the Y gray scale images are used for feature extraction and classification.

In order to test the ability of the algorithm to against noise, salt and pepper noise is added to the images, and edge detection is carried out in ZMs, The calculation of ZMs follows Eq. (3 ~ 6) and Eq. (11 ~ 12), where the edge weight l_t is set to 0.11 and K_T is determined to be 24. Sobel method, LoG method, SUSAN method, FIR method, MMG method and other edge



detection methods are used for comparison. The edge detection results of each method are shown in Fig. 8.

It can be seen from the Fig. 8 that: 1) the methods of Zernike, Sobel and LoG are unaffected by additive noise. However, the edge of the image processed by Sobel method has obvious discontinuity, which indicates that the edge information is lost more. LoG is too sensitive and the edge features contain too much redundant information. Correspondingly, the edge features extracted by Zernike method have higher accuracy and less information loss. Theoretical studies show that Zernike moments have higher accuracy in image edge location than Sobel; 2) FIR and MMG can extract image edge features, but the robustness against noisy is not ideal, moreover, the accuracy of the edge extracted by MMG method is low, and the image edges are overlapped, which is difficult to distinguish; 3) SUSAN failed to extract the edge features of the image effectively.

After data preprocessing, the images used to input the networks have obvious contour features, which provide a good condition for the CNNs to classify. Next, the images will be classified by improved CNNs.