

## Research Article

# Real-Time Identification of Nonstandard Thermal Inkjet Codes on the Surface of Hot-Rolled Steel Sheets in Complex Environments

**Xing-Hua Wang** , **Le Liang**, **Hong-Wei Sun**, **Wei-Bin Chen**, **Can-Can Wang**,  
and **Yang-Yang Sun**

*Jiangsu Automation Research Institution, Lianyungang, Jiangsu 222006, China*

Correspondence should be addressed to Xing-Hua Wang; wangxinghua2009@sina.cn

Received 19 April 2022; Revised 24 May 2022; Accepted 30 May 2022; Published 14 July 2022

Academic Editor: A. Parthiban

Copyright © 2022 Xing-Hua Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Many steel mills currently rely on thermal spray codes to carry and convey information about the steel plate at different stages of production. Thermal inkjet codes are patterns of numbers, letters, and symbols formed by continuous spraying of high-temperature resistant coatings on the surface of billets, so it is especially important to identify thermal inkjet code information accurately and timely during the processing stage of steel plates. Most of the existing mainstream identification methods rely on manual visual observation and traditional image recognition to obtain them, with low recognition efficiency and poor recognition effect. In this paper, a lightweight YOLO with ResNet18 as the backbone network is used as the detection and recognition framework, and deformable convolution and text feature extraction techniques are purposefully added to detect the location of thermal inkjet codes images to achieve accurate and fast positioning and segmentation of thermal spray codes. Meanwhile, an attention mechanism-based character recognition module is added to quickly infer the content within the ROI location of thermal inkjet codes. The experimental results show that the thermal inkjet code character recognition method proposed in this paper has a high recognition rate and fast recognition speed, which meets the practical application requirements of relevant enterprises.

## 1. Introduction

The production of hot-rolled steel sheets has many stages, and errors in data positioning flow [1] not only cause damage to production efficiency but also, if similar production data are replaced by different steel sheets and shipped out, may cause quality objections from customers and lead to compensation losses from manufacturers or induce safety production accidents. For the above problems caused by steel plate flow data, they are usually circumvented by strengthening manual responsibility training, multilayer personnel confirmation, and factory random inspection [2, 3], and more advanced methods such as steel plate production digital twin with automatic product identification in the form of manual confirmation [4]. But these solutions have the following drawbacks: (1) high uncertainty, the process of adding or subtracting steel plates is prone to errors in the flow of data, resulting in unnecessary production losses; (2) strong

manual subjectivity, due to the decline in the quality of the thermal spray code resulting in the wrong image recognition, manual modification is easy to mis-enter; and (3) low efficiency, manual observation, as well as the modification, is time-consuming and prone to cause blockage in the flow of steel plates. With the development of technology, there are also some new technologies applied to the data information flow of hot-rolled steel plates, such as RIFD technology [5], laser coding [6], and so on. With these technologies, the real-time flow of steel plate data information can be realized, but the reliability of the equipment and the efficiency of information flow has not achieved significant improvement, but on the contrary, new production and manpower input problems have been introduced.

With the rapid development of neural network technology and continuous iteration, the integration of image processing technology and neural networks provides the possibility of real-time and accurate identification of

nonstandard hot-rolled steel plate surface thermal spray codes. Neural networks, especially convolutional neural networks (CNNs), have been widely used in tasks such as detection, recognition, and segmentation due to their good generalization ability and parameter-sharing properties. Currently, the detection algorithms based on convolutional neural networks can be generally classified structurally as two-stage algorithms represented by faster R-CNN [7] and one-stage algorithms represented by YOLO [8, 9].

In this paper, a combination of detection and recognition is used for the real-time recognition of nonstandard thermal inkjet codes. First, a lightweight YOLO based on the residual network ResNet18 is used as the backbone network, and deformable convolution is added as a text region detection method to achieve region localization and segmentation of multiscale features, followed by an irregular text feature extractor and an attention mechanism-based recognition network to complete the text inference task, and finally achieve real-time and accurate thermal inkjet code recognition.

The YOLO (You Only Look Once) series network is a regression-based one-stage target detection algorithm, which integrates classification and border regression into a vector inside as the direct output of the result, so the YOLO algorithm detection is better in real time. The modified lightweight YOLO network in this paper consists of four residual modules, and the ResNet18 network is used as the backbone network for feature extraction. Each residual module consists of multiple small residual units connected sequentially. Each residual module consists of two DCBL modules and one residual edge. Each DCBL module consists of three components: deformable convolutional layer (DCN), batch normalization layer (BN), and activation function (Leaky ReLU), and each component is combined with the residual connection method.

For efficiency, a lightweight residual network, ResNet18, is used as the backbone network. Obviously, lightweight backbone networks tend to produce features with smaller acceptance domains and weaker representation capabilities. To address this issue, we use the construction of an improved YOLO detection network by adding deformable convolution (DCN). Conventional convolutional networks use square convolution with regular sampling points that limit their ability to model geometric transformations and can only scan with fixed shape and size windows. In contrast, each element of the deformable convolution kernel [10, 11] has a learnable parameter bias, which allows the sampling points of the deformable convolution to be adaptively adjusted according to the feature map, and the perceptual field can change with the shape and size of the object.

In the subsequent text feature inference process of YOLO, the seq2seq model based on multiheaded attention [12] is used as the recognition module in this paper. The module consists of an initiator instead of an encoder, together with a decoder to form a recognition network. The initiator is used to find the starting position of any text string (SOS). An embedding layer  $E1$  with linear transformations and a multiattentive layer  $A1$  together form the initiator. The decoder consists of two LSTM layers [13] and a multiheaded

attention layer  $A2$ . The recognition module has a simple structure and maintains good accuracy for thermal inkjet character recognition.

## 2. Experimental Details

*2.1. General Framework.* In order to adapt the application requirements of real time and accuracy of nonstandard thermal inkjet code recognition in hot-rolled steel production scenarios, a deep learning method for nonstandard thermal inkjet code detection and recognition based on improved YOLO detection network and recognition network based on attention mechanism is with the overall framework, as shown in Figure 1, which mainly consists of four parts: lightweight backbone network, pyramid enhancement module, inkjet code detector, and recognizer based on the attention mechanism.

### *2.2. Improved Lightweight YOLO Framework for Thermal Inkjet Code Detection and Recognition Algorithm*

*2.2.1. Improved YOLO Model.* For target detection in complex environments, deep-learning-based detection algorithms are excellent. Compared with target detection frameworks such as RCNN, the YOLO model has higher accuracy and detection speed. Considering that the specific application is the area image detection of thermal inkjet code, which is a single-category multitarget detection with a simpler task, and in order to reduce the computational effort for faster processing, the classical ResNet18 with fewer layers is chosen as the backbone network; while combining deformable convolution and the seq2seq model based on the attention mechanism, an improved lightweight YOLO model is proposed, which contains the following four improvement ideas:

*(1) Deformable Convolution.* In this paper, based on the YOLO feature extraction backbone network based on the ResNet18 residual network, we construct an improved YOLO detection network by adding deformable convolution (DCN). The sampling points of conventional convolutional network rules can only be scanned with a window of fixed shape and size. In contrast, each element of the deformable convolution kernel has a learnable parameter bias, which allows the sampling points of deformable convolution to be adaptively adjusted according to the feature map, and the perceptual field can change with the shape and size of the object.

Conventional convolution is divided into two steps: (1) sampling by a fixed-size convolution kernel on the input feature map and (2) computational sampling based on the convolution kernel weights and summing of the computational results. The convolution kernel of  $3 \times 3$  and expansion rate of 1 are defined as  $R$ :

$$R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}. \quad (1)$$

Then, for a point  $p_0$  on the output feature map, the regular convolution can be expressed as follows:

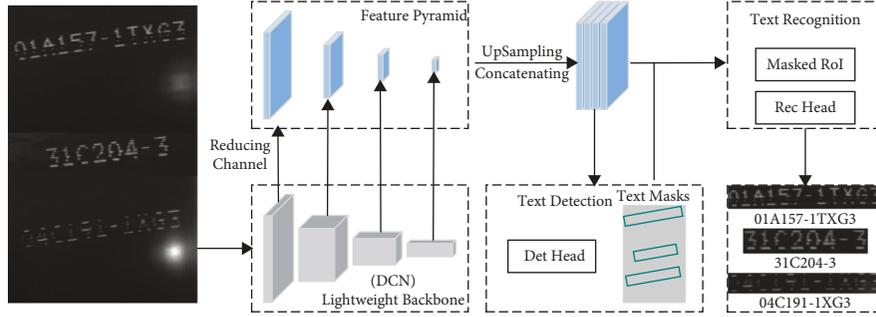


FIGURE 1: Overall architecture of nonstandard thermal injection code detection and identification.

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n), \quad (2)$$

where  $w$  is the convolution kernel weight and  $x$  is the input feature map.

The deformable convolution adds a two-dimensional bias  $\Delta p_n$  to each sample point position of the conventional convolution kernel and predicts a weight  $\Delta m_n$  for each sample point, as shown in Figure 2. At this point, for a point  $p_0$  on the output feature map, the deformable convolution is computed as follows:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n). \quad (3)$$

Since the offsets  $\Delta p_n$  are usually fractional, the value of  $x$  needs to be calculated by bilinear interpolation as follows:

$$x(p) = \sum_q G(q, p) \cdot x(q), \quad (4)$$

where  $p$  denotes the arbitrary position after the offset, that is,  $p = p_0 + p_n + \Delta p_n$ ;  $x(q)$  is the pixel value at the four adjacent integer coordinates around  $p$ , and  $G$  is the weight corresponding to the four integer coordinates, which can be divided into two one-dimensional kernels.

$$G(q, p) = g(q_x, p_x) \cdot g(q_y, p_y), \quad (5)$$

where  $g(m, n) = \max(0, 1 - |m - n|)$ .

(2) *Irregular Text Feature Extractor*. For distorted thermal inkjet images, this paper proposes an ROI extractor based on the literature [14] for extracting fixed-size feature blocks of distorted thermal inkjet text lines. First, the minimum outer rectangle of the thermal inkjet text is calculated; second, the text feature region inside the rectangle is obtained; then, the noise interference is significantly reduced by multiplying the text feature region with a binary mask with all weights of 0 outside the thermal inkjet text; finally, the feature block is adjusted to a fixed size as follows:

$$F_{roi} = \text{Resize}_{w \times h}(\text{Crop}(F_f, \text{Rect}(m_i)) \circ m_i), \quad (6)$$

where  $m_i$  – the binary mask for the first line of the thermal inkjet code,  $\text{Rect}(\cdot)$  – the smallest outer rectangle of the binary mask,  $\text{Crop}(F_f, \cdot)$  – the operation of cropping feature patches in the vertical ROI region,  $F_f \circ$  – the element

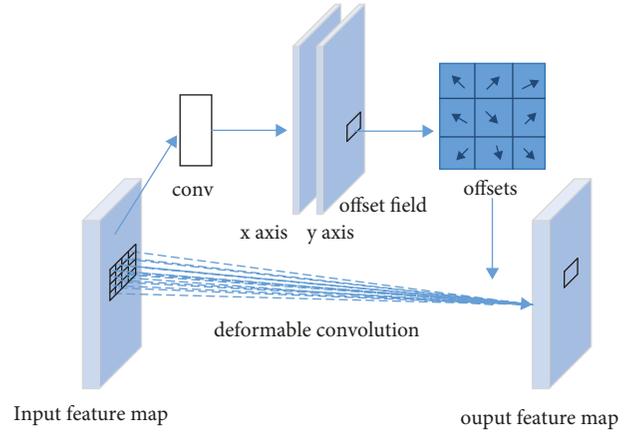


FIGURE 2: Deformable convolution.

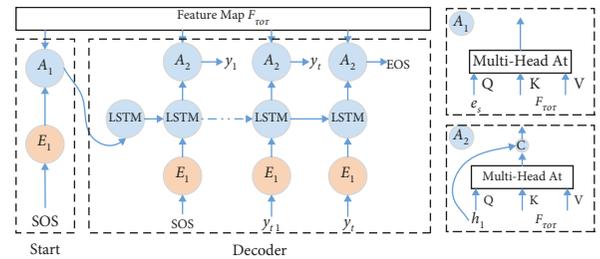


FIGURE 3: Details of the recognition head.

multiplication, and  $\text{Resize}_{w \times h}(\cdot)$  – the operation to adjust the feature map to  $w \times h$  by bilinear interpolation.

(3) *Attention Mechanism*. In this paper, the recognition head is composed of several attention heads of the seq2seq model. It is shown in Figure 3. It consists of an initiator and a decoder. The initiator consists of an embedding layer  $E_1$  using a linear transformation and a multiheaded attention layer  $A_1$ , both with the number of layers being one. As shown in equation (7). The embedding layer  $E_1$  converts the SOS symbol as a hot into a 128-dimensional vector, after which the 128-dimensional vector enters the multiheaded attention layer together with the feature block  $F_{roi}$  to generate the SOS feature vector, the input vector for the initial time-step decoder.

$$f_s = A_1(E_1(\text{"SOS"}), F_{roi}). \quad (7)$$

Two LSTM layers and a multiheaded attention layer form a decoder. At the initial time step, the feature vector  $f_s$  of SOS and the zero LSTM initial state are used as the input to the decoder. After that, the hidden state  $h_0$  of the initial time step is input to the LSTM along with the SOS symbols, and the output symbol  $y_1$  is predicted at time step 1. Finally, the output symbol  $y_1$  is used as the input to the LSTM, and the loop ends at the prediction string (EOS) marker. The process is as in the following equation:

$$\begin{aligned} y_t &= \operatorname{argmax}(\operatorname{FC}(A_2(h_t, F_{roi}))), 1 \leq t \leq T, \\ h_0 &= \operatorname{LSTM}(f_s, 0), \\ h_t &= \operatorname{LSTM}(E_2(y_{t-1}), h_{t-1}). \end{aligned} \quad (8)$$

(4) *Loss Function.* In order to achieve the positioning and recognition of nonstandard thermal spray codes, the loss function of this paper is as follows:

$$L = L_{\text{det}} + L_{\text{rec}}, \quad (9)$$

where  $L_{\text{det}}$  represents the nonstandard heat code detection loss function and  $L_{\text{rec}}$  represents the identified loss function.

$$L_{\text{det}} = L_{\text{tex}} + \alpha L_{\text{ker}} + \beta(L_{\text{agg}} + L_{\text{dis}}), \quad (10)$$

where  $L_{\text{tex}}$  is the loss of thermal coding area segmentation,  $L_{\text{ker}}$  is the loss of thermal coding character segmentation, and  $\alpha$  and  $\beta$  are used to adjust the weights of  $L_{\text{tex}}$ ,  $L_{\text{ker}}$ ,  $L_{\text{agg}}$ , and  $L_{\text{dis}}$ . In all experiments,  $\alpha = 0.55$  and  $\beta = 0.3$ .

$$\begin{aligned} L_{\text{tex}} &= 1 - \frac{2 \sum P_{\text{tex}}(i)G_{\text{tex}}(i)}{\sum P_{\text{tex}}(i)^2 + \sum G_{\text{tex}}(i)^2} \\ L_{\text{ker}} &= 1 - \frac{2 \sum P_{\text{ker}}(i)G_{\text{ker}}(i)}{\sum P_{\text{ker}}(i)^2 + \sum G_{\text{ker}}(i)^2}, \end{aligned} \quad (11)$$

where  $P_{\text{tex}}(i)$  and  $G_{\text{tex}}(i)$  denote the value of the  $i$ -th pixel and the ground truth value of the inkjet code region in the segmentation result, respectively.  $P_{\text{ker}}(i)$  and  $G_{\text{ker}}(i)$  denote the  $i$ -th pixel value and the ground truth value in the thermal inkjet character prediction, respectively.

The loss function  $L_{\text{rec}}$  for character recognition is given by

$$L_{\text{rec}} = \frac{1}{|\omega|} \sum_{i=0}^{|\omega|} \operatorname{CrossEntropy}(y_i, \omega_i), \quad (12)$$

where  $\omega$  is the ground transcript containing the EOS symbol,  $\omega$  is the amount of characters in the inkjet code process, and  $\omega_i$  is the first character in the inkjet code.

2.3. *Thermal Spray Code Detection Dataset Construction.* Due to the complex scene characteristics of the actual production of hot-rolled steel plates, the thermal spray code images of steel plates captured by the camera are often affected by some environmental factors, resulting in the brightness or darkness of the image. Therefore, it is necessary

to preprocess the actual captured global image of the thermal inkjet code, and this paper uses contrast enhancement to preprocess the captured image of the steel plate surface containing the thermal inkjet code to improve the quality of the original thermal inkjet image features. Furthermore, considering that the deep network model requires a very large amount of label data, the actual nonstandard thermal spray code image data with different defects and problems that can be collected in the steel production site are limited to meet the training requirements of the deep learning model. Therefore, this paper uses a mosaic-based data augmentation approach to achieve an effective increase in the amount of thermal spray code image data [15]. The framework includes various ways of image expansion, such as elastic change, rotation, Gaussian noise, and HSV offset. Based on the above four data enhancement algorithms, the collected thermal code images of steel plate production sites are enhanced and constructed to obtain thermal code detection and recognition data sets for the training and learning of lightweight YOLO models.

### 3. Results and Discussion

3.1. *Dataset Construction.* Considering that deep learning models require a large number of labeled samples to complete training and learning, this paper uses mosaic-based data enhancement to expand the actual collected thermal code images, including elastic change, rotation, HSV offset, and Gaussian noise, and then constructs the hot-rolled steel plate code detection and recognition data set.

In this paper, the nonstandard thermal inkjet images captured on the surface of hot-rolled steel plates are enhanced by mosaic-based data enhancement, using elastic change, rotation, HSV offset, and added noise to expand the actual captured thermal inkjet images by 250, 350, 400, and 700 images, respectively, which in turn expands the original thermal inkjet image data set from 300 to 2,000 images and the original thermal inkjet symbol data set from 2,369 to 16,000. The expanded thermal inkjet data set is shown in Table 1.

Similar to the literature [16, 17], the nonstandard thermal inkjet code data set constructed based on Table 1 and in the ratio of 9:1 is used to produce the training set and test set required for the lightweight YOLO model proposed in this paper. Among them, the total number of data sets used for thermal inkjet image detection is 2,000, 1,800 for the training set, and 200 for the test set; the total number of data sets used for character detection is 16,000, 14,400 for the training set, and 1,600 for the test set.

3.2. *Experimental Environment and Parameter Settings.* To ensure the training efficiency of the lightweight YOLO network model, the experimental environment configuration shown in Table 2 is used for the training and testing of the whole model. The operating system is Ubuntu 16.0.4; the CPU is Intel Core i7-7820X, the memory is 16 GB; the graphics card is NVIDIA GeForce RTX 2080Ti (11 GB); and the training platform of the model is torch 1.6.0, the deep

TABLE 1: Image data expansion results.

	Thermal code image	Thermal code character	Total
Original data set	300	2,369	2,669
Extensional data set	2,000	16,000	18,000

TABLE 2: Experimental environment.

Category	Parameter
System	Ubuntu 16.0.4
CPU	Intel core i7-7820X CPU @ 3.60 GHz
GPU	GeForce RTX 2080Ti 11G
RAM	16G
Relations	CUDA = 10.1, OPENCV, python = 4.4.0, torch = 1.6.0

learning development tool, and the programming language is Python.

In addition, the proposed lightweight YOLO model was first pretrained in Image Net [15], after which the training parameters for thermal spray code detection recognition were set as follows: learning rate of 0.001, batch size of 4, and Epoch set to 240 times.

**3.3. Performance Evaluation Criteria.** To measure the performance of the network model, this paper uses performance evaluation metrics commonly used for target detection, including average precision (*AP*), mean average precision (*mAP*), precision rate (*P*), recall rate (*R*), and F1 value. *P* is the ratio of the number of correct predictions by the model for a class to the number of all predictions for that class and is calculated as follows:

$$R = \frac{T_P}{T_P + F_N}, \quad (13)$$

where *R* is the ratio of the number of correct detections of a category of the model to the total number of that category in the test set and is calculated as follows:

$$P = \frac{T_P}{T_P + F_N}. \quad (14)$$

The *F1* value is a weighted summed average of precision and recall and is a common evaluation metric for a comprehensive assessment of performance in classification problems, as follows:

$$F1 = 2 \times \frac{RP}{R + P}. \quad (15)$$

*AP* is the area of the precision-recall curve, and the higher the *AP* value of a category, the better the detection of that category. *mAP* is a comprehensive evaluation index.

### 3.4. Experimental Results and Comparison

**3.4.1. Image Contrast Enhancement.** The actual nonstandard thermal inkjet image is enhanced with the gamma transform-based contrast operation in Section 1, and the image

effect after the image enhancement process is shown in Figure 4. Figure 4(a) is the original image, and Figure 4(b) is the image obtained after image contrast enhancement. We can see that the contrast of the thermal coding image is significantly enhanced and the image feature information is more obvious.

**3.4.2. Preselection Box Selection and Setting.** As the non-standard thermal inkjet codes and character targets are more fixed, the effect of using the original preselection box size of YOLO is poor, and the preselection box size needs to be reset. However, since the type and target size of the non-standard thermal inkjet code data set are relatively single, the clustering centers obtained directly by the *k*-means clustering algorithm tend to be more concentrated, which affects the overall model training effect.

For this reason, an improved preselected frame size clustering method is used in this paper. First, the *k*-means method is used to cluster the manually labeled nonstandard thermal inkjet data set to obtain the set preselected box size. Based on this, the preselected box size is further stretched by linear scale scaling operation to obtain the preselected box size that meets the needs of nonstandard thermal inkjet code detection and recognition of hot-rolled steel sheets.

$$\begin{cases} (w'_1, h'_1) = 0.5(w_1, h_1), \\ (w'_9, h'_9) = 3(w_9, h_9), \\ (w'_i, h'_i) = \left( \frac{(w_i - w_1)}{w_9 - w_1} (w'_9 - w'_1), w'_i \frac{h_i}{w_i} \right), \end{cases} \quad (16)$$

where *w* and *h* represent the width and height of the *k*-means clustered preselected box size, respectively. *w*<sub>1</sub> and *h*<sub>1</sub> represent the width and height of the preselected box size after linear size scaling, respectively. The preselected boxes are sorted according to their sizes. *i* is the number of the clustered preselected box,  $1 \leq i \leq 9$ . Therefore, the corresponding preselected box sizes of nonstandard thermal inkjet codes for hot-rolled steel sheets obtained by the above calculation are shown in Table 3. Table 3 shows the results of *k*-means preselected box clustering.

**3.4.3. Model Training Parameters' Setting.** To demonstrate the training process of the improved lightweight YOLO model more visually, the learning curves of the model are given in Figure 5. Red is the learning curve of the thermal coding detection model, and blue is the learning curve of the code character recognition model. When the number of iterations exceeds 240, the whole model tends to converge



FIGURE 4: Image preprocessing effect: (a) original thermal inkjet coding image and (b) enhanced image results.

TABLE 3: *K*-means clustering results of prior box.

Feature layer	7 × 7	14 × 14	28 × 28
Thermal coding detection model	(104, 59)	(112, 74)	(45, 13)
	(121, 68)	(129, 127)	(91, 38)
Character recognition model	(208, 64)	(194, 40)	(101, 32)
	(156, 190)	(95, 128)	(39, 25)
	(163, 217)	(126, 166)	(44, 67)
	(197, 276)	(141, 233)	(90, 108)

TABLE 4: Comparison of different deep learning detection models.

Network	mAP (%)	FPS
Faster-RCNN	95.76	12
YOLO-tiny	97.11	38
YOLO-v3	99.58	25
Lightweight YOLO (CNN)	95.58	55
Lightweight YOLO (DCN)	99.54	54

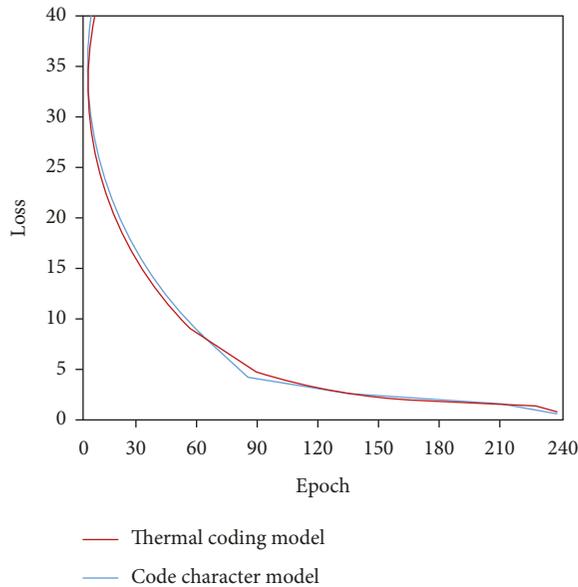


FIGURE 5: Learning curve of lightweight YOLO model.

stably. Therefore, the number of iterations of Epoch is set to 240 in the experiment.

**3.4.4. Lightweight YOLO Model Performance Analysis.** The size of the improved lightweight YOLO network model is only 10 MB, which significantly reduces the device performance requirements for subsequent deployment. Based on the training and test sets in Section 2.1, four models, faster-RCNN, YOLO-tiny, YOLO-v3, and the conventional convolutional network model, were selected as the comparison models in order to highlight the advantages of the improved lightweight YOLO model in terms of detection accuracy and real-time performance, and the performance comparison results of different models are given in Table 4. Table 5 gives the results of nonstandard

TABLE 5: Lightweight YOLO test results.

Category	Precision (%)	Recall (%)	F1
0	99.78	99.67	0.997
1	99.23	99.01	0.991
2	97.52	99.25	0.984
3	99.12	100.00	0.996
4	100.00	99.98	1.000
5	98.87	100.00	0.994
6	99.24	99.28	0.993
7	99.49	99.89	0.997
8	100.00	100.00	1.00
9	98.70	100.00	0.993
Thermal coding detection	99.54	100.00	0.998
Character recognition location	98.87	100.00	0.994

thermal inkjet detection and character localization for the improved lightweight YOLO model. Figure 6 gives the results of nonstandard thermal inkjet code detection with character recognition.

According to the experimental results shown in Tables 4 and 5, for the needs of nonstandard thermal inkjet code detection and recognition of hot-rolled steel sheets with real-time performance, the proposed algorithm of nonstandard thermal inkjet code detection and recognition of hot-rolled steel sheet surface based on the improved YOLO model of ResNet [18] has better detection and recognition accuracy and robustness compared with traditional image processing methods and other deep network target detection algorithms.

Specifically, compared with the traditional YOLO-v3 model, the accuracy mAP of the lightweight YOLO model is greatly improved from the original 24 frames per second to 54 frames per second with only a 0.04% reduction in the detection inference speed FPS. Compared with the conventional convolutional network (CNN), the accuracy mAP of the model with deformable convolution (DCN) is improved by 3.96%. Compared with the other two models (faster-RCNN and YOLO-tiny), the detection accuracy of the proposed lightweight YOLO model is improved by



FIGURE 6: Detection and identification results of nonstandard thermal inkjet code.

3.78% and 2.43%, respectively, while the detection inference speed FPS is improved by 42 frames per second and 16 frames per second, respectively. The experimental results illustrate the effectiveness of the proposed improved method and also prove that the proposed lightweight YOLO model can obtain better detection accuracy and higher inference speed, which is more in line with the practical application requirements of fast detection and identification of hot-rolled steel surface thermal inkjet codes. In addition, from the recognition results of each category, the accuracy of the lightweight YOLO model designed with the improved recognition module reaches more than 98% in each category and still has better accuracy and robustness in the presence of strong noise or other complex situations in the test images. All the above experimental results prove the effectiveness and reliability of the proposed nonstandard thermal code detection and recognition algorithm for hot-rolled steel sheets based on the lightweight YOLO model.

#### 4. Conclusion

For the practical application requirements of hot-rolled steel surface thermal inkjet code detection and recognition, this paper proposes a nonstandard thermal inkjet code detection and recognition method based on the lightweight YOLO model of ResNet18. The method is based on the YOLO model, and through the design of an improved recognition module, the YOLO backbone feature extraction network is designed to be lightweight, and then a lightweight YOLO model is obtained, which effectively reduces the computational and storage complexity of the model with almost no loss of accuracy. In addition, this paper collects thermal inkjet code data from actual hot-rolled steel enterprises under production scenarios and uses the mosaic framework

for data enhancement to construct a nonstandard thermal inkjet code detection and recognition data set on the surface of hot-rolled steel sheets. The experimental results show that the proposed nonstandard thermal inkjet code detection and recognition method based on the ResNet18 lightweight YOLO model can achieve real-time and accurate thermal inkjet code detection and character localization and recognition. The field performance test results show that the algorithm can fully meet the efficiency and accuracy requirements of nonstandard thermal inkjet code data flow on the surface of the high-speed production site of steel plates and has been practically applied in the actual production process of hot-rolled steel plates. The algorithm can be further optimized later to improve the generalization ability of the algorithm while ensuring the detection accuracy and recognition speed. Also, how to detect small probability errors in time is the focus of future research.

#### Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

#### Conflicts of Interest

The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

#### Acknowledgments

This work was supported by Jiangsu Key R & D Plan (BE2020105) and the National Key Research and Development Program of China (no. 2018YFB1305900).

## References

- [1] Y. Chao, Y. Xiao, and Q. Huang, "Analysis of situation and development tendency of hot strip production," *Mechanical Engineering and Automation*, vol. 12, no. 6, pp. 124–126, 2005.
- [2] L. H. Ying, W. D. Xun, and C. T. you, "Study of manufacturing execution system on iron and steel enterprise," *Metallurgical Industry Automation*, no. 5, pp. 34–39, 2003.
- [3] Z. Zheng, K. Zhang, and X. Gao, "Human-cyber-physical system for production and operation decision optimization in smart steel plants," *Science China Technological Sciences*, vol. 65, no. 2, pp. 247–260, 2022.
- [4] Z. j. Xu, Z. Zheng, and X. q. Gao, "Operation optimization of the steel manufacturing process: a brief review," *International Journal of Minerals Metallurgy and Materials*, vol. 28, no. 8, pp. 1274–1287, 2021.
- [5] J. Zhou, D. Niu, Q. Li, and J. Liu, "Iron and steel ladles tracking management system based on RFID and WLAN," *Journal of Engineering*, vol. 2019, no. 22, pp. 8310–8314, 2019.
- [6] L. Lyubomir, T. Edmunds, K. Tsanko, and A. Nikolay, "Influence of power density and frequency of the process of laser marking of steel products," *Infrared Physics & Technology*, vol. 56, 2021.
- [7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [8] R. Joseph, S. Kumar Divvala, and B. Ross, "Girshick, ali farhadi. you only look once: unified, real-time object detection," *CoRR*, 2015.
- [9] Y. Ding, S. Zhou, H. Yuan et al., "Crack identification method of steel fiber reinforced concrete based on deep learning: a comparative study and shared crack database," *Advances in Materials Science and Engineering*, vol. 68, 2021.
- [10] F. Chen, F. Wu, J. Xu, G. w. Gao, G. Qi, and X. Yuan, "Adaptive deformable convolutional network," *Neurocomputing*, vol. 56, 2020.
- [11] X. Cheng and C. Zhen zhong, "Multiple video frame interpolation via enhanced deformable separable convolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, 2021.
- [12] Z. Zhou and F. Liu, "Filter gate network based on multi-head attention for aspect-level sentiment classification," *Neurocomputing*, vol. 55, pp. 214–225, 2021.
- [13] G. Liu and J. Guo, "Bidirectional LSTM with attention mechanism and convolutional layer for text classification," *Neurocomputing*, pp. 325–338, 2019.
- [14] S. Y. Qin, A. Bissacco, M. Raptis, Y. Fujii, and Y. Xiao, "Towards unconstrained end-to-end text spotting," *CoRR*, vol. 909231 pages, 2019.
- [15] T. Fujisawa and K. Saitoh, "Bayesian direct-binary-search algorithm for the efficient design of mosaic-based power splitters," *OSA Continuum*, vol. 4, no. 4, p. 1258, 2021.
- [16] De Y. Li, G. Fa Wang, Y. Zhang, and S. Wang, "Coal gangue detection and recognition algorithm based on deformable convolution YOLOv3," *IET Image Processing*, vol. 16, no. 1, pp. 134–144, 2021.
- [17] K. J. Wu, C. Bai, D. Wang, Z. Liu, T. Huang, and H. Zheng, "Improved object detection algorithm of YOLOv3 remote sensing image," *IEEE Access*, vol. 9, pp. 113889–113900, 2021.
- [18] L. Fei Fei, J. Deng, and K. Li, "ImageNet: constructing a large-scale image database," *Journal of Vision*, vol. 9, no. 8, p. 1037, 2010.