H. REN, Y. J. ZHANG, J. T. CHEN, X. N. WEI, H. K. CHEN, P. LIU

ISSN 0543-5846
METABK 63(3-4) 423-426 (2024)
UDC – UDK 669.14-621-41:621.771-536.45:620.191.001.5=111

A METHOD FOR DETECTING SURFACE DEFECTS IN HOT-ROLLED STRIP STEEL BASED ON DEEP LEARNING

Received – Primljeno: 2024-02-04
Accepted – Prihvaćeno: 2024-02-30
Original Scientific Paper – Izvorni znanstveni rad

Hot-rolled strip steel is a material widely used in production activities and daily life. However, the appearance of surface defects during its production process is inevitable. To address this issue, we introduce a new detection method using Gold-Yolo to detect surface defects on hot-rolled strip steel. Our method effectively balances accuracy and real-time performance while detecting four common types of surface defects, achieving an average accuracy rate of 82,2 % for detecting individual types of surface defects. Experimental data prove that our method excels in classifying and locating surface defects on hot-rolled steel strip, demonstrating broad application prospects and promotional value.

Keywords: steel strip, hot-rolled, surface defect, object detection, Gold-Yolo

INTRODUCTION

Hot-rolled strip steel is a key product in the steel industry, produced through the hot rolling process. This process involves heating steel billets to high temperatures (usually above 1 000 °C) and then rolling them in a continuous mill set in their hot state to achieve the desired thickness and width. This manufacturing method endows the steel with good plasticity and ductility, making it suitable for subsequent processing and applications. It is widely used in construction, the automotive industry, shipbuilding, and machinery manufacturing, among other fields. However, even minor surface defects can lead to serious structural problems, especially in applications where there are strict requirements for material strength and durability. Therefore, defect detection in modern industry is necessary.

In the past, the detection of surface defects on steel plates mainly relied on manual methods, which had significant subjectivity and could not be performed in real-time. In recent years, with the development of machine vision and image processing technologies, most manufacturers have begun to adopt these technologies. Current steel plate surface defect detection techniques based on machine vision and Convolutional Neural Networks (CNN) have achieved good results. However, traditional machine vision technologies have some shortcomings, such as poor adaptability.

In view of this, we propose a single-stage steel plate surface defect detection method based on Gold-Yolo. This method introduces new features and improvements while retaining the basic framework, enabling it to classify defects and determine their locations, thereby significantly increasing the detection speed. Compared to traditional methods, Gold-Yolo offers higher efficiency and precision, especially when dealing with complex or irregular surfaces. Through this method, we can more effectively identify and locate various defects on steel plates, thereby ensuring product quality and improving production efficiency.

RALATED WORK

In the development of surface defect detection and classification methods for hot-rolled steel strips, early methods relied on manual summarization of image features and used matching methods to detect defects. For example, literature [1] proposed an "anti-noise method based on complete local binary patterns" and constructed a dataset based on NEU-DET. However, manual summarization and matching of defect images require domain experts to analyze images in specific environments to obtain more precise features.

Subsequent research, such as in literature [2], proposed a method based on CNN (Convolutional Neural Networks) for detecting surface defects on hot-rolled steel strips. Literature [3] introduced a method based on Random Forest and Support Vector Machines. Compared to traditional machine learning methods, deep neural networks are more prominent in the field of automatic image feature extraction, with CNNs being widely applied. Literature [4] proposed a defect detection algorithm combining Swin Transformer and a multi-threshold structure.

To effectively meet the high-performance requirements during the production of hot-rolled strip steel,

H.Ren,Y.J.Zhang,J.T.Chen,X.N.Wei,H.K.Chen,P.Liu, School of Computer Science and Software Engineering, University of Science and Technology Liaoning, China. Corresponding author: Y. J. Zhang (1997zyj@163.com)

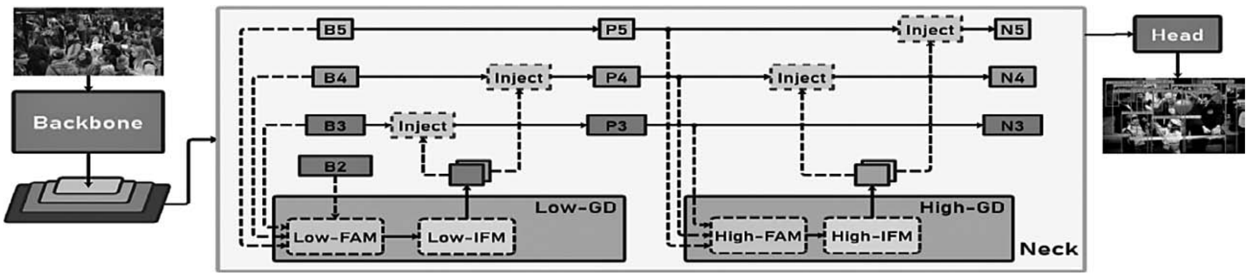

Figure 1 The network structure of Gold-Yolo

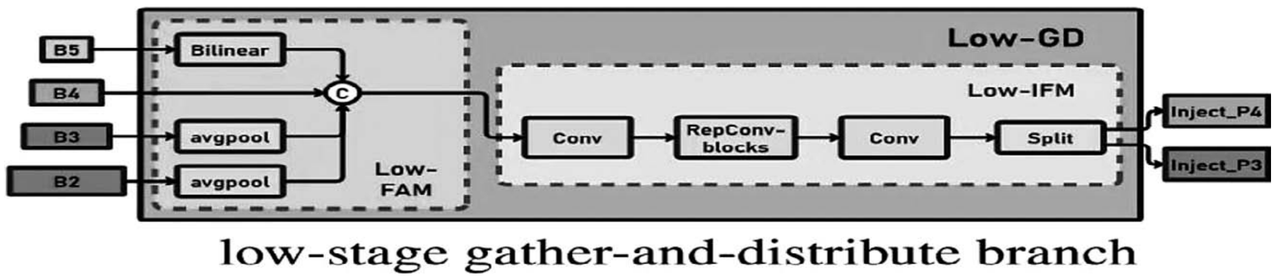

Figure 2 Gather-and-Distribute structure

this article adopts the latest Gold-Yolo model. Gold-Yolo introduces the "Gather-and-Distribute" (GD) mechanism, effectively overcoming the limitations of traditional FPN (Feature Pyramid Network) structures in information transmission and enhancing the detection capability for objects of different sizes. Specifically, Gold-Yolo's Low-Order Gather and Distribute (Low-GD) and High-Order Gather and Distribute (High-GD) branches can efficiently process defect features of varying sizes. The low-order branch focuses on capturing small-sized defects, while the high-order branch is more suited for detecting large-sized defects. In this way, Gold-Yolo maintains both high accuracy and speed, meeting the strict real-time performance requirements in the production of hot-rolled steel strips.

METHODOLOGY

The network structure of Gold-Yolo is as shown in Figure 1, it is very similar to the traditional Yolo model, but it differs in the Neck stage.

The Yolo series utilizes the classic FPN (Feature Pyramid Network) architecture in its intermediate layers. This architecture achieves multi-scale feature fusion through multiple branches. However, this structure mainly optimizes the integration of features between adjacent layers, relying on an indirect, recursive method for information fusion at more distant levels.

In the traditional FPN structure, there is a significant issue in information transmission: a considerable amount of information may be lost during the transfer process. This is because the interaction between layers is primarily limited to the information selected by the intermediate layers, with unselected information being lost during transfer. As a result, the information from a specific level mainly provides effective support only to its adjacent layer, offering limited assistance to the broader global hierarchy. Therefore, overall, the effectiveness of information fusion may be somewhat restricted.

Gold-Yolo has improved the traditional Feature Pyramid Network (FPN) structure by adopting a novel "Gather-and-Distribute" (GD) mechanism. In conventional FPN structures, the fusion of features from different levels often leads to information loss. This is especially true during cross-level information exchange, where the transmission of information is not direct enough, resulting in inefficiency and information loss. Gold-Yolo addresses this issue with its innovative GD mechanism, which uses a unified module to collect and fuse information from different levels, and then effectively distributes this information to each level. This not only prevents information loss during transmission but also enhances the information fusion capability without significantly increasing latency.

The GD mechanism of Gold-Yolo includes three main modules: the Feature Alignment Module (FAM), the Information Fusion Module (IFM), and the Information Injection Module (Inject). The FAM is responsible for collecting and aligning features from different levels, while the IFM fuses these features to generate global information. Finally, the Inject module distributes this information to each level, enhancing the detection capability of the branches through simple attention operations. Additionally, Gold-Yolo has developed two branches: the Low-GD branch and the High-GD branch. These branches extract and fuse features from large and small-sized feature maps, respectively, to enhance the model's detection capabilities for objects of different sizes.

In Figure 2, the Low-Order Stage Branch features the Low-Order Feature Alignment Module (Low-FAM) and the Low-Order Information Fusion Module (Low-IFM), collectively known as the Low-Order Gather and

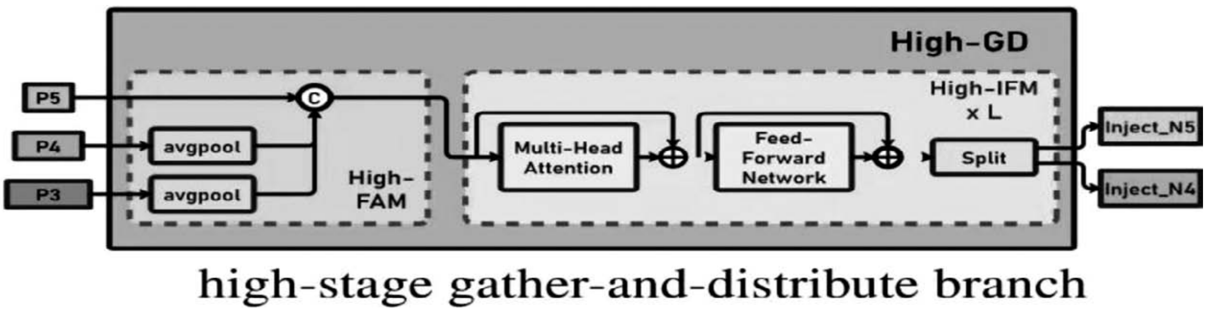

Figure 3 Gather-and-Distribute structure

Distribute Branch (Low-GD). This branch focuses on merging output features from the Backbone network (e.g., B2, B3, B4, B5) to obtain high-resolution features that retain information about small objects. Within this branch, the Low-Order Feature Alignment Module (Low-FAM) uses an average pooling (AvgPool) operation to downsample input features and achieve a uniform size. Then, information fusion is carried out through a Transformer module to minimize computational complexity.

In Figure 3, there is the High-Order Feature Alignment Module (High-FAM) and the High-Order Information Fusion Module (High-IFM), together forming the High-Order Gather and Distribute Branch (High-GD). This branch integrates features generated by the

Low-GD (such as P3, P4, P5). The High-Order Feature Alignment Module (High-FAM) includes the use of average pooling (avgpool) to reduce the dimensions of input features to a uniform size. The High-Order Information Fusion Module employs multiple stacked Transformer blocks, each containing multi-head attention blocks, a feed-forward network (FFN), and residual connections.

METHOD OF IMPLEMENTATION

Prior to training the Gold-Yolo model, preprocessing was conducted for the four categories in the NEU-DET dataset, which included data cleaning and image augmentation. The aim was to expand the dataset and enhance the model's generalization performance. The dataset was then divided into a training set and a test set in an 8:2 ratio. The model's trainable parameters were initialized randomly. The training set, after being batch-processed, was used for training the model. Adjustments to the model were made through backpropagation. After each training epoch, the model's trainable parameters were fixed, and the test set was used to evaluate the model's performance. This process continued until there was no further improvement in the model's detection performance.

EXPERIMENT AND ANALYSIS

For our study, we selected four types of defects from the NEU-DET steel surface defect dataset: Inclusions, Patches, Pitted Surface, and Scratches. These defects were manually annotated in the dataset, with the corresponding results illustrated in Figure 4.

During the training process of the Gold Yolo model, we referred to the training configuration and network architecture of Yolov6-3,0 (except for the middle layer part) for setup. The model's backbone network utilized the EfficientRep Backbone, while the detection head part employed an efficient decoupled head structure. To optimize the training process, we adopted the Stochastic Gradient Descent (SGD) method with momentum and applied a cosine decay strategy for adjusting the learning rate. Additionally, the model training included a warm-up phase, grouped weight decay, and Exponential Moving Average (EMA). In terms of training strategy, we also incorporated self-distillation and Anchor-based Assistant Training (AAT). As for data augmentation techniques, weused methods like Mosaic and Mix-up. After 150 epochs of training, the learning rate was reduced to 0,01. The training was conducted on an RTX 3050 device and lasted a total of 1,5 hours.

$$\rho interp\left(r_{n+1}\right)=\max_{\tilde{r}:\tilde{r}\geq r_{n+1}}\rho\left(\tilde{r}\right) \tag{1}$$

$$AP=\sum\nolimits_{r=0}^{1}\left(r_{n+1}-r_{n}\right)\rho interp\left(r_{n+1}\right) \tag{2}$$

$$mAP=\frac{\sum\nolimits_{i=1}^{k}AP_{i}}{k} \tag{3}$$

Where *pinterp* (r_{n+1}) is the interpolated precision at a specific recall level r_{n+1}. It is calculated by taking the maximum precision $\rho(\tilde{r})$ for all recall levels $\tilde{r}$ that are greater than or equal to r_{n+1}, $\rho(\tilde{r})$. This is the precision at the recall level $\tilde{r}$, AP Average Precision, which is computed as the sum of the product of the interpolated precision *pinterp* (r_{n+1}) and the difference in recall levels ($r_{n+1}-r_n$) over all recall levels from 0 to 1, r_n and r_{n+1}. These are the adjacent recall levels. When evaluating the Gold Yolo model, we used the Mean Average Precision (mAP), which considers both Precision and Recall. Precision refers to the proportion of correctly predicted positive samples among all samples predicted as positive. Recall represents the proportion of correctly predicted positive samples among all actual positive samples. The calculation of Average Precision (AP) involves averaging the precision at each recall level.

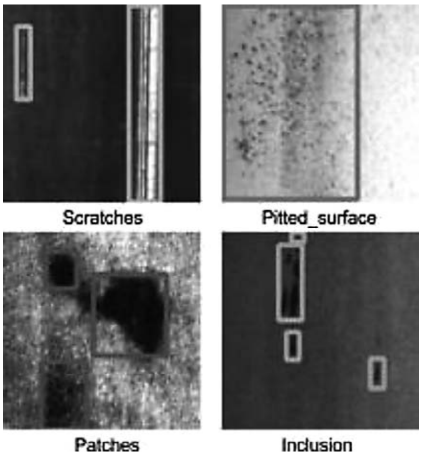

Figure 4 Annotation sample by Labelimg

These precision values are based on the ascending order of recall values and are derived using formulas (1) and (2). Finally, the mAP is calculated using formula (3), which averages the AP values for each category, where k represents the number of categories.

CONCLUSION

The researcher selected 20 % of the images from the training set that did not participate in the training as the test set. On the NEU-DET dataset, the average accuracy of the Gold Yolo model reached 82,2 %, while the average accuracy using Yolov8 under the same conditions was only 81 %. The processing speed of this method on an RTX 3050 graphics card was 33,3 frames per second, with each image's analysis time being only 30 milliseconds. This speed is sufficient to meet the requirements of surface defect detection and localization during the high-speed operation of hot-rolled steel strip lines.

Acknowledgements

This work was supported in part by innovation and entrepreneurship training program for college students of University of science and Technology Liaoning 2024.

REFERENCESf

[1] A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNetClassification with Deep Convolutional Neural Networks. Neural Information Processing Systems (NIPS), New York, 2012, 1097-1105

[2] K Peng, X Zhang. Classification technology for automatic surface defects detection of steel strip based on improved BP algorithm Fifth International Conference on Natural Computation. IEEE, (2009) 1,110-114.

[3] E Amid, S R Aghdam, H Amindavar. Enhanced per- formance for support vector machines as multiclass cla- ssifiers in steel surface defect detection. International Journal of Electrical and Computer Engineering 6(2012) 7,693-697.

[4] B. Yu, X. K. Zhang, and W. Wang. "Surface Defect Detection of Hot-Rolled Strip Based on STMR-CNN." Computer Systems & Applications 31.10 (2022), 122-133.

Note: The responsible translators for English language is J. Wang – University of Science and Technology Liaoning, China