Intelligent Machine Vision for Detection of Steel Surface Defects with Deep Learning

Yi Huang

July 8, 2023

Intelligent Machine Vision for Detection of Steel Surface Defects with Deep Learning

Yi Huang
Department of Mechanical Engineering
National University of Singapore
Singapore, Singapore
eastonhwang@gmail.com

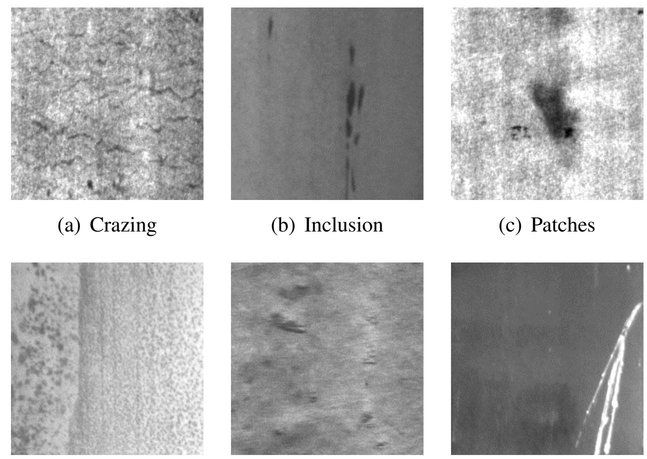
***Abstract*—With the high demand for steel surface quality, the requirement for defect-free steel surfaces is growing. Recent applications of deep learning in machine vision have demonstrated impressive performance. Our work aims to look into efficient surface defects detection algorithms, and to attempt to improve defect detection performance. This paper reports the use of YOLOv5 for steel surface defect detection and achieving 95.9% mean average precision(mAP). Moreover, we have improved detection accuracy by preprocessing the database with filters and denoisers based on CNNs. After applying denoisers and filters, apparent improvement can be seen in each type of defect after using either one of the techniques. For example, after applying denoisers and filters, the detection average precision(AP) of Rolled-in Scale defects increased by 12.6% and 35.4%, respectively. In this paper, the efficiency of machine vision based on deep learning, and the effectiveness of preprocessing in improving accuracy for steel surface defect detection are demonstrated.**

***Index Terms*—Machine Vision, Defect Detection, Deep Learning, Denoising**

I. Introduction

With the development in industry, such as aerospace and shipping, there is a growing demand for higher steel surface quality. However, the manual inspection process not only was insufficient to guarantee the defect-free surface of steel products, but was also time-consuming. Recent applications of deep learning in machine vision have demonstrated impressive performance. In this work, we aim to look into efficient algorithms for detecting steel surface defects. Faster R-CNN [1] and YOLO series [2] are two typical and widely-applied detection algorithms so far. The difference between them is that YOLO regards object detection as a regression problem, while Faster R-CNN divides object detection into two parts, classification and location. YOLO detects objects faster, but the accuracy is less than Faster R-CNN. However, Jeong-ah Kim reported that the accuracy of YOLOv3 is higher than Faster R-CNN in vehicle recognition [3].

In this paper, we attempted to apply the YOLO series for steel surface detection and to improve detection accuracy through pre-processing of the database with denoisers and filters. We used the database from Northeast University (NEU) [4], which contains 1800 images, including six types of hot rolled steel strip surface defects: crazing (Cr), inclusion (In), patches (Pa), pitted surface (PS), rolled-in scale (RS), and scratches (Sc). Fig. 1 shows samples of each type of defect.



(a) Crazing (b) Inclusion (c) Patches
(d) Pitted Surface (e) Rolled-in Scale (f) Scraches

Fig. 1. Samples of Each Type of Defects

II. Literature Review

Deep learning is a common method of target detection. The deep learning network has an important impact on the accuracy and speed of target detection. Kaiming proposed a residual learning framework(Resnet) to ease the training of substantially deeper networks than those used previously [5]. Simonyan K proposed a thorough evaluation of networks(VGGnet) of increasing depth using an architecture with tiny convolution filters [6]. Szegedy C proposed inception V3 with a small amount of computation, but high performance is offered, 2.5 times higher than v1 [7]. At present, many researchers are studying target detection. Redmon J proposed that YOLO uses the whole graph as the network's input and directly returns the location and category of the bounding box in the output layer [8]. Bochkovskiy A proposed YOLOv4, which improves the accuracy by applying some methods to enhance CNNs [9]. There is also many papers on the extraction and identification of surface defects. Changsheng Li proposed a new method of defect extraction for mobile phone screens based on machine vision [10]. Melanie Po-Leen Ooi put forward a defect extraction scheme for defect-cluster identification [11]. Piervincenzo Rizzo improved the general guided-wave technique with magnetostrictive transducers for the detection and sizing of defects in strands [12]. Generally,

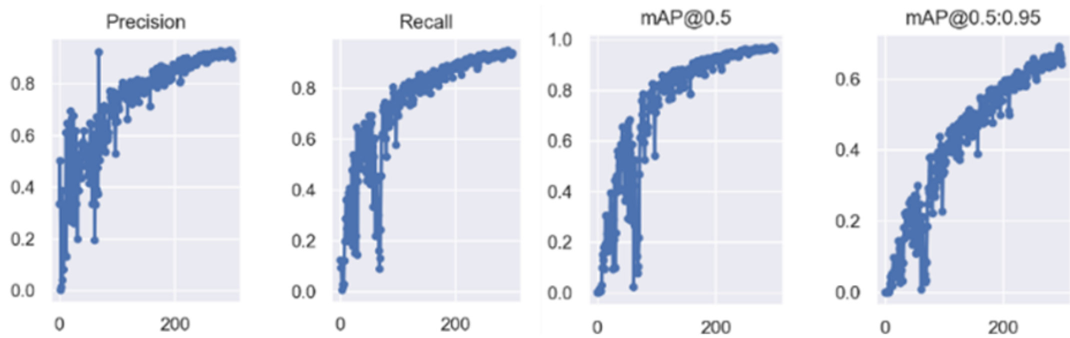


Fig. 2. The Indicators during The Training Process

noise has a significant impact on industrial image quality. Therefore, image noise reduction significantly improves defect recognition accuracy. Soh J W proposed a universal blind denoiser(DUBD), which can reduce noise from various real environments [13]. Kai Zhang presented a denoising convolutional neural network, FFDNet, to achieve fast denoising [14]. Kai Zhang proposed DnCNN to handle Gaussian denoising with the unknown noise level [15].

III. Defects Detection

In this paper, we first applied YOLOv5 [16], the latest version of YOLO series to detect steel surface defects. YOLO frames object detection as a regression problem to spatially separated bounding boxes and associated class probabilities. YOLO can divide images into a grid system, while each cell in the grid is responsible for detecting objects within itself.

We set the image size to 224 pixels, epochs to 300, batch size to 128 images, and trained the YOLOv5s model with GPU. The database was divided into training and testing sets and contained 1500 and 300 images, respectively.

A. Training Evaluation and Detection Result

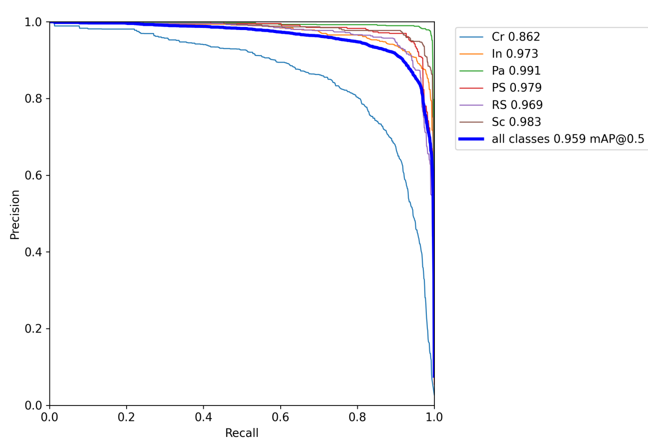


Fig. 3. Precision - Recall Curve

1) Evaluation Indicators: In detection evaluation, there are four basic indicators, which are the combinations of truth(true or false) and prediction(positive or negative), including True Positive(TP), True Negative(TN), False Positive(FP), and False Negative(FN). With these four basic indicators, we can calculate precision(Eq. 1), recall(Eq. 1) and F_1 score(Eq. 2), as follows:

$$Precision = \frac{TP}{TP+FP}, \quad Recall = \frac{TP}{TP+FN}, \tag{1}$$

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{2}$$

2) Training Evaluation: Fig. 2 presents the record of some essential indicators during the training process. The value of these indicators fluctuated greatly in the early training, but these indicators increased stably later and approached the value of "1" finally. However, since they were still growing around "1", more epochs are needed for further training. Precision (P) and recall (R) are contradictory when detecting defects. When one increases, the other decreases, and it can be seen in Fig. 3. However, we can evaluate the performance from the area enclosed by the P-R curve (Fig. 3). According to the P-R curve, the crazing(Cr) defects detection performs the worst, while patches(Pa) defects performs the best. The same results can be seen in F_1 score curve (Fig. 4). Furthermore, since F_1 balances between P and R, the F_1 curve provides the best value of the threshold of confidence, at point (0.462, 0.92).

TABLE I
Detection Evaluation

mAP	**AP**					
	Cr	In	Pa	PS	RS	Sc
0.959	0.842	0.967	0.988	0.995	0.987	0.976

3) Detection Result: Table I illustrates the evaluation indicators of the model, including mean average precision (mAP) at 0.959 for all types of defect detection and average precision (AP) of each defect. According to the table, the model has an accurate detection capability, but for the images with a lot of noise, especially Cr, the model performed not satisfactorily.

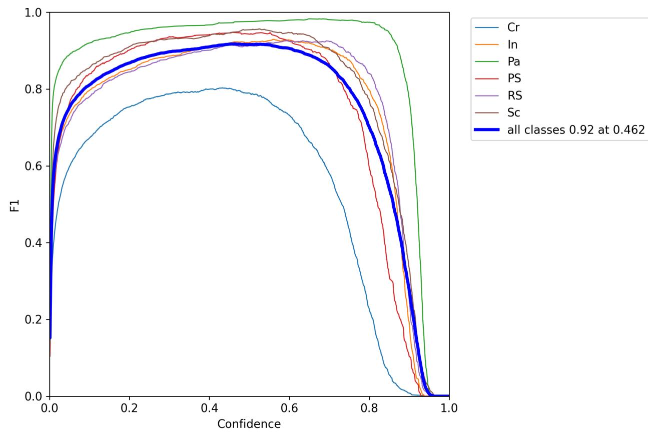
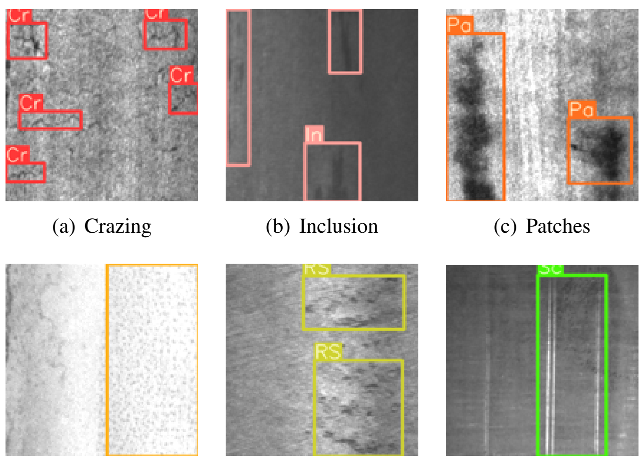


Fig. 4. F_1 Score - Confidence Curve



(a) Crazing (b) Inclusion (c) Patches

(d) Pitted Surface (e) Rolled-in Scale (f) Scraches

Fig. 5. Samples of Defects Detection Result

The detection results are shown in Fig. 5. The detection platform can accurately classify and locate each type of defect, but there is still many areas to improve the accuracy with the pre-processing the database.

IV. Pre-processing the Database

Due to the harsh industrial manufacturing environment, industrial images inevitably contain unnumbered noises, which can be seen through grayscale value analysis, shown in Fig. 6. From the analysis, the noise blurs the defects. Moreover, the noise will change the defects' characteristics and affect the performance of defect detection.

A. Denoisers based on CNNs

Denoising techniques, based on deep learning, are increasingly popular since their strong learning ability and effectiveness. To improve detection accuracy, we have applied several denoisers based on the various CNNs, including DRUNet [17], FFDNet, SRMD [18], IRCNN [19], Esrgan [20], DUBD and SwinIR [21]. Table II shows the evaluation results of DRUNet and DUBD that performed well among those CNNs.

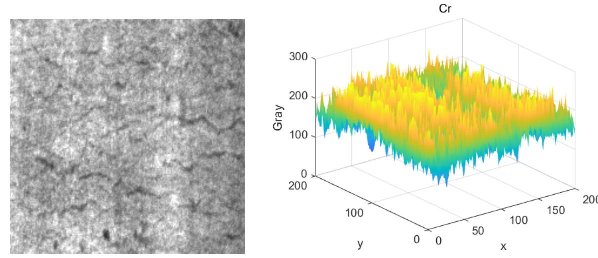


Fig. 6. Grayscale Value Analysis

TABLE II
Performance Comparison of Different Combinations of Denoisers

Training Set Denoiser	Testing Set Denoiser	mAP	AP					
			Cr	In	Pa	PS	RS	Sc
Original	Original	0.607	0.373	0.313	0.908	0.8	0.413	0.832
	DRUNet	0.546	0.246	0.271	0.89	0.497	**0.539**	**0.834**
	DUBD	0.57	0.305	0.293	0.891	0.581	0.519	0.83
DRUNet	Original	0.555	0.409	0.334	0.918	**0.859**	0.151	0.658
	DRUNet	0.56	0.306	0.275	0.925	0.625	0.394	0.837
	DUBD	0.582	0.341	0.263	**0.941**	0.784	0.328	**0.834**
DUBD	Original	**0.61**	**0.439**	**0.347**	0.899	**0.859**	0.307	0.807
	DRUNet	0.577	0.335	0.281	0.925	0.607	0.485	0.828
	DUBD	0.593	0.362	0.296	0.932	0.693	0.47	0.807

1) Method and Parameters: In the denoisers(DRUNet and DUBD), we used the pre-trained models provided in [17 & 13] to denoise training(TR) and testing(TE) set. Then we used the TR sets after preprocessing to train models with 50 epochs using YOLOv5s, and obtained two new models (with DRUNet and DUBD as denoisers) besides the original model. Finally, we use these models to test different TE sets and evaluated their performance of defect detection.

2) Results Analysis: Comparing with the original combination(TR and TE), models with the database after denoising have better detection performance as shown in Table II. The improvement is evident in the single type of defect detection, especially in the images polluted by the noise greatly. For example, the average precision(AP) of using TR with DUBD denoiser and TE with original increased by 6.6% in detecting crazing(Cr) defects. As for roll-in scale(RS) defects detection, the combination of using TR with original and TE with DRUNet increased by 12.6%, which is a vast improvement. Preprocessing by denoisers shows its effectiveness in accuracy improvement of steel defect detection.

B. Filters

Filters were used widely in smoothing the noise. In this section, Median Filter (MF) and Gaussian Filter (GF) were applied to smooth the database.

1) Method and Parameters: Same as the denoisers section's experiment steps. Filters' kernel size was set to 5 and trained the TR by YOLOv5s for 50 epochs.

2) Results Analysis: The detection evaluations are illustrated in Table III. From the result, the filter combinations are superior to the original combination. The best combination of TR and TE by GF, improves the mAP by 22%(0.827 vs

0.607). The immense improvement can be seen in the single type of defects. For example, the average precision(AP) of crazing(Cr) and rolled-in scale(RS) detections are increased by more than 20%. Especially, the AP of RS detection with database preprocessed by GF is 35.4%(0.767 vs 0.413) more than the original combination. In addition, the majority of the best scores of AP are conducted on the combination of training and testing sets by GF. The GF is most effective in smoothing this database so far for this work-in-progress studies.

TABLE III
PERFORMANCE COMPARISON OF DIFFERENT COMBINATIONS OF FILTERS

Training Set Filter	Testing Set Filter	mAP	AP					
			Cr	In	Pa	PS	RS	Sc
Original	Original	0.607	0.373	0.313	0.908	0.8	0.413	0.832
	Gaussian	0.687	0.466	0.818	0.89	0.682	0.341	0.925
	Median	0.544	0.215	0.781	0.848	0.268	0.255	0.895
Gaussian	Original	0.569	0.357	0.483	0.839	**0.862**	0.112	0.761
	Gaussian	**0.827**	**0.622**	**0.866**	**0.963**	0.823	**0.767**	0.922
	Median	0.794	0.567	0.842	0.949	0.802	0.674	0.933
Median	Original	0.511	0.331	0.394	0.781	0.821	0.086	0.656
	Gaussian	0.766	0.524	0.843	0.928	0.772	0.607	0.919
	Median	0.814	0.588	0.843	0.968	0.797	0.754	**0.933**

V. CONCLUSIONS

In this paper, we have detected all types of defects based on YOLOv5 with 95.9% mAP and almost every kind of defect detection AP up to more than 95%, except for the crazing defects, which has a lot of noise in the images. Furthermore, we have demonstrated the effectiveness of preprocessing the database by denoisers and filters to improve defect detection accuracy. This improvement can be seen in every type of defect. Especially, denoisers and filters improved the performance of detecting the crazing defects, and increased the average precision by 6.6% and 24.9%. We have found that denoising data effectively improves detection accuracy.

We will continue to work on this project and make improvements in future work, such as use of other target algorithms besides YOLOv5, improve the parameters of the denoiser, and add a priori condition to the target detection with the physical characteristics of defects.

REFERENCES

[1] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28.

[2] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.

[3] Kim J, Sung J Y, Park S. Comparison of Faster-RCNN, YOLO, and SSD for real-time vehicle type recognition[C]//2020 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia). IEEE, 2020: 1-4.

[4] Song K, Hu S, Yan Y. Automatic recognition of surface defects on hot-rolled steel strip using scattering convolution network[J]. J. Comput. Inf. Syst, 2014, 10(7): 3049-3055.

[5] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[6] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale visual recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

[7] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2818-2826.

[8] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.K. Elissa, "Title of paper if known," unpublished.

[9] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.

[10] Li C, Zhang X, Huang Y, et al. A novel algorithm for defect extraction and classification of mobile phone screen based on machine vision[J]. Computers & Industrial Engineering, 2020, 146: 106530.

[11] Ooi M P L, Joo E K J, Kuang Y C, et al. Getting more from the semiconductor test: Data mining with defect-cluster extraction[J]. IEEE Transactions on Instrumentation and Measurement, 2011, 60(10): 3300-3317.

[12] Rizzo P, di Scalea F L. Feature extraction for defect detection in strands by guided ultrasonic waves[J]. Structural Health Monitoring, 2006, 5(3): 297-308.

[13] Soh J W, Cho N I. Deep Universal Blind Image Denoising[C]//2020 25th International Conference on Pattern Recognition (ICPR). IEEE, 2021: 747-754.

[14] Zhang K, Zuo W, Zhang L. FFDNet: Toward a fast and flexible solution for CNN-based image denoising[J]. IEEE Transactions on Image Processing, 2018, 27(9): 4608-4622.

[15] Zhang K, Zuo W, Chen Y, et al. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising[J]. IEEE transactions on image processing, 2017, 26(7): 3142-3155.

[16] Jocher, G., Nishimura, K., Mineeva, T., Vilariño, R.: YOLOv5 (2020). https://github.com/ultralytics/yolov5. Accessed 10 July 2020

[17] Zhang K, Li Y, Zuo W, et al. Plug-and-play image restoration with deep denoiser prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.

[18] Zhang K, Zuo W, Zhang L. Learning a single convolutional super-resolution network for multiple degradations[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 3262-3271.

[19] Zhang K, Zuo W, Gu S, et al. Learning deep CNN denoiser prior for image restoration[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 3929-3938.

[20] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4681-4690.

[21] Liang J, Cao J, Sun G, et al. Swinir: Image restoration using swin transformer[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1833-1844.