An Efficient Industrial Product Serial Number Recognition Framework

Mitchel M. Hsu, Ming-Hsien Wu, Yun-Chieh Cheng, Chia-Yu Lin
Department of Computer Science and Engineering
Yuan Ze University, Taoyuan, Taiwan
Email: sallylin0121@saturn.yzu.edu.tw

2022 IEEE International Conference on Consumer Electronics - Taiwan | 978-1-6654-7050-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCE-TAIWAN55306.2022.9869266

***Abstract*—The serial number is essential for product tracking in industries. Without serial numbers, the control of production is lost. Optical character recognition (OCR)—a technique for extracting characters' features and utilizing these features to recognize them in digital images—is widely used to identify serial numbers and improve product tracking efficiency. However, it is relatively challenging for OCR to adapt to digital images containing unclean environments in industries. Images taken in such environments can include various objects, making it hard for OCR models to focus on serial numbers. In this paper, we propose an industrial product serial number recognition framework, which can efficiently adapt complex structures of the background environment and recognize serial numbers in industries. The framework includes the data preprocessing stage, detection stage, and recognition stage. Eventually, our framework reaches nearly four times better on the testing dataset than the baseline score and increases the efficiency of product tracing in industries.**

I. Introduction

With improving production efficiency, recording and tracing product numbers have a vital role in the industry. However, manually recording these numbers can put a heavy burden on employees, leading to decreasing efficiency. Currently, more and more firms are adopting OCR for serial number recognizing and tracking. At the same time, there are various practical approaches to improve the accuracy of optical character recognition, including adjusting the identification evaluation indicators, which enable granularity and completeness as a basis for calculating identification scores [1]. Furthermore, [2] and [3] uses various neural network architectures and OCR algorithms as the model structure for the processing flow to increase the accuracy and overall flexibility in recognizing text. In addition, it is also a common practice to change the way of data preprocessing and noise cancellation according to different data properties, such as scene text [4], industry serial number [5]. Despite the recent superb performance of OCR in various fields, we also need to be aware that the results can be gaping differences when a well-established processing pipeline and a robust neural network correspond to different situations. Therefore, the emphasis in this article will not only be on the industrial application of OCR but also express some concepts in the process flow to provide an efficient processing framework to deal with issues that need to be considered when doing OCR. The framework includes three stages: the data preprocessing stage, the detection stage, and the recognition stage. The data preprocessing stage utilizes median filters to remove noise from images. The detection stage includes serial number locating and reverse serial number classification. Serial number locating detects serial number by Yolov5 and crop it down from industrial products. The recognition stage uses Transformation-ResNet-None-Attention (TRNA) model to recognize the serial number in the image. The main contributions of this paper are as follows:

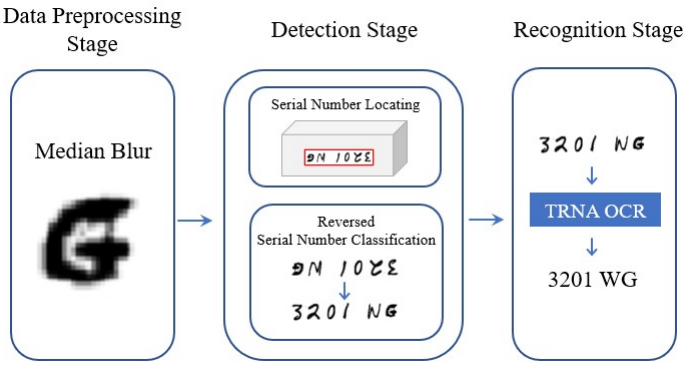

Fig. 1. An efficient industrial product serial number recognition framework.

- Represent a framework that can detect the serial number from products in complex industrial environments.
- Help mitigate employees' workload and increase the efficiency of product tracing.

II. Methods

In this research, we propose an efficient framework for recognizing industrial product serial numbers. The framework includes the data preprocessing stage, the detection stage and the recognition stage, as shown in Fig. 1.

A. Data Preprocessing Stage

We are not allowed to show the data we adopted in this paper due to the privacy policy. However, we will demonstrate resemblance data constructed with the MNIST dataset [6] and the EMNIST dataset [7]. Serial numbers on industrial products are mostly handwritten with paint guns, as shown in Fig. 2. Handwritten texts may lead to burry characters as well as connected brush strokes, such as the character W in Fig. 2. In order to alleviate the impact of these problematic handwritten features on our models, we first apply a median filter to our

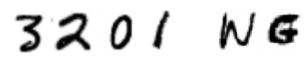

Fig. 2. Handwritten serial number constructed with the MNIST dataset and the EMNIST dataset.

training data. Median filters can remove noise from images by blurring them blur unnecessary strokes.

B. Detection Stage

In the detection stage, two main tasks are addressed: locating the serial numbers on products and reversing serial numbers on up-side-down products.

1) Locating Serial Numbers: Product serial numbers are usually captured with large and complex backgrounds that make serial numbers too tiny to be identified. In order to remove backgrounds in our images and zoom into the serial numbers, we adopted and trained Yolov5, a state-of-the-art object detecting model, to accomplish this task. However, it is possible to have multiple products in a single image, yet our target is to identify the serial number on the closest product in each image. Thus, we only crop down the serial number with the largest bounding box, which will be the one on the closest product in the corresponding image.

2) Reversing Up-Side-Down Cases: Industrial products have a high chance to be placed up-side-down on the conveyor leading to reverse their serial numbers as well. Reversed text can cause OCR models to confuse when recognizing them, such as the character M and W. To resolve this problem, we built and trained an image classification model transferred from VGG16 using the cropped images in the last part. This model is to identify serial numbers that are up-side-down, and justify them into the correct position.

C. Recognition Stage

In the final part of our framework, the task is to recognize the serial number in the image we processed in the last two stages using an OCR model. In this research, we adopted the TRBA model [3], one of the best models competing in ICDAR challenges. The model includes four stages of image processing, namely the transformation stage, feature extraction stage, sequence modeling stage, and prediction stage. However, our task is to recognize serial numbers on industrial products, which are usually not sequential. As a result, we removed the sequence modeling stage in the TRBA model to be our OCR model named TRNA.

III. Experiment

We used a dataset including a training set with 12067 images and a testing set with 6037 images in this research. The baseline model that only uses the TRBA scored 335 on Levenshtein distance on the testing set. In this section, the results of our experiment are discussed stage by stage. In the data preprocessing stage, we adopted a median filter using a 5×5 kernel size to mitigate connected strokes for handwritten characters. Then, in the detection stage, serial numbers are extracted and justified using the Yolov5 object detection model and a reversed image classification model. The object detection model reached 0.98 on mAP.5 on the testing set, and the reversed image classification model reached 0.99 on accuracy on the testing set. Finally, the extracted serial number is recognized using the TRNA OCR model. It reached 93 on Levenshtein distance with the testing set, nearly four times better than the baseline model.

IV. Conclusion

In this research we presented an efficient framework for recognizing serial numbers on industrial products. Our framework includes three stages: the data preprocessing stage, the detection stage, and the recognition stage. In the data preprocessing stage, we successfully reduced the influence of inconsistency in handwritten characters by applying a median filter to the dataset. Furthermore, in the detection stage, we overcome complex environments in the industry and boosted the performance of the OCR model when performing recognition tasks by using Yolov5 as a serial number detector and a reversed character classifier transferred from VGG16. Finally, we adopted one of the finest OCR models competed in the ICDAR challenges in the recognition stage. However, we removed the sequence modeling stage from this model due to the non-sequential feature of industrial serial numbers. For experiments, the serial number detector can achieve over 0.96 on mAP.5, and the reverse character classifier can achieve over 0.99 on accuracy. Nonetheless, the OCR model we trained can reach 93 on Levenshtein distance, nearly four times better than the baseline model. As a result, the proposed framework achieved a high accuracy towards serial number detection and recognition. This study provides a more efficient way for industrial product tracking by representing this framework. Nevertheless, boosting industry's production efficiency.

References

[1] Lee, Chae Young, Youngmin Baek, and Hwalsuk Lee. "Tedeval: A fair evaluation metric for scene text detectors." 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW). Vol. 7. IEEE, 2019.

[2] L. Yu, "State-of-the-art OCR Algorithm for Translation Guiding: An Online-offline Combinational Model," 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), 2021, pp. 213-217, doi: 10.1109/ICIRCA51532.2021.9544563.

[3] Baek, Jeonghun, et al. "What is wrong with scene text recognition model comparisons? dataset and model analysis." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.

[4] K. S. Satwashil and V. R. Pawar, "Integrated natural scene text localization and recognition," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), 2017, pp. 371-374, doi: 10.1109/ICECA.2017.8203708.

[5] Choi, Eunjeong, Somi Chae, and Jeongtae Kim. "Machine learning-based fast banknote serial number recognition using knowledge distillation and Bayesian optimization." Sensors19.19 (2019): 4218.

[6] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 86(11):2278-2324, November 1998.

[7] Cohen, G., Afshar, S., Tapson, J., and van Schaik, A. (2017). EMNIST: an extension of MNIST to handwritten letters. Retrieved from http://arxiv.org/abs/1702.05373.