Neural network component of the product marking recognition system on the production line

Kroshchanka Aliaksandr
Brest State Technical University
Brest, Belarus
kroschenko@gmail.com

Ivaniuk Dzmitry
JSC "Savushkin product"
Brest, Belarus
id@pda.savushkin.by

***Abstract*—The paper considers the implementation of an intelligent computer vision component based on a neural network approach to solve the problem of recognizing various product markings manufactured by JSC Savushkin Product. A feature of the system is the use of modular architecture, which makes it easy to add new models. The proposed system is a component of a more general neurosymbolic system.**

***Keywords*—Computer vision, ANN, deep neural networks, IIoT, object detection**

I. Introduction

The competitiveness of any production is based on the release of various and high-quality products. Quality is the defining criterion when a client chooses a particular product, and diversity allows to cover various groups of potential buyers. Systems that automate the quality control process while maintaining the variety of products in a large enterprise are of particular value. At the same time, the quality control process is carried out not only for the product itself, but also for the packaging and marking that the buyer sees. Since the packaging forms the initial impression of the product, its quality is one of the reasons that gives the buyer a reason to buy. Marking as an element of the packaging guarantees the buyer the safety of the product for the specified period, subject to the storage conditions.

Recently, the methods of artificial intelligence in general and machine learning in particular have been widely used in industrial systems installed in enterprises to control the production process. Intelligent subsystems simplify many of the routine operations carried out to maintain the quality of the finished product. For example, control over the correctness of marking was previously carried out exclusively by a human operator. Now, with the evolution of the theory of computer vision, which has received a significant breakthrough thanks to ideas from the constantly evolving field of learning deep neural networks, systems are being created that allow this operation to be performed faster and more regularly than a human could do.

In addition to tracking markings of only one human-readable type (for example, alphanumeric), the modular implementation of such systems makes it possible to universalize the recognition process and easily add prepared models that recognize new types of markings (for example, a Data Matrix code). This special type of matrix barcode allows you to encode special identification information, as well as weight, expiration date, serial number, consignment and date of manufacture of the product [1].

In addition to universality in the processing of products of various types, automation of control in production is of particular importance. With the evolution of the concept of Industry 4.0 and IIoT (Industrial internet of things), more and more intelligent control solutions are being created. The implementation of such solutions allows full control over production, including control over the condition of production equipment and safety measures. Moreover, these important functions are carried out without human intervention. In the future, this can lead to the creation of completely autonomous production facilities, in which even the operations of planning the purchase of raw materials and the shipment of finished products will be carried out automatically.

This work is devoted to the development of a neural network component that is part of a more general neurosymbolic system described in [2].

II. Problem formulation

The task set before us was to develop a system for verifying the correctness of marking on products manufactured by JSC Savushkin Product. An important aspect is that the system must work in real time, based on data from a camera installed above the production line. This camera generates a video stream at 76 frames per second.

In addition to the alphanumeric code (Fig. 1), since recently, products can be produced with alternative marking including a Data Matrix code (Fig. 2) [3]. This type of marking is a convenient and concise representation of the specific and basic data about the product.

Figure 1. Product with alphanumeric marking

Figure 2. Product with Data Matrix code

Based on the general formulation of the problem, the following subtasks can be distinguished, which must be solved by the system:

1) detection and recognition of the type of marking;
2) marking recognition;
3) identifying possible marking problems.

In general, the following requirements can be imposed on the system that will solve the assigned tasks:

- **High speed of work**. The production line is moving very quickly, so the detection of defects should be carried out with minimal delays;
- **Autonomy**. The system should minimize operator involvement in the quality control process;
- **Universality**. The system should be configured to recognize the marking of any product;
- **Adaptability**. The system must work under any conditions that occur in production (for example, insufficient lighting, personnel errors, etc.).

Let's list the main problems with marking [2]:

1) **lack of ink:** in case of receipt of products on the production line without marking, the system must conclude that there is no ink;
2) **camera shift:** if no data on the recognition results are received from the neural network modules, but the system knows that movement along the production line has begun, then it must conclude that the camera has shifted;
3) **bad marking:** marking was found and recognized, but did not match the template representation. In this case, the system must conclude that the marking is incorrect;
4) **unreadable marking:** if the marking is blurry and cannot be recognized, it is necessary to stop the production line and report the error to the operator.

For problems 1,3 and 4, it is necessary to screen out products that have a problem marking. The occurrence of these problems implies a complete stop of the production line movement and reporting to the operator about the problem.

III. Overview of existing approaches

Despite the existing interest in the autonomy of production processes and the indisputable advantages that its implementation entails, tasks similar to those described are solved in a large number of cases with the participation of a person. The operator simply checks a part of the production periodically and randomly. This approach has disadvantages:

- only a small part of the production passes inspection, so there is a possibility that the defective marking will be missed;
- the speed of a person's reaction to an emergency situation may be insufficient;
- a person may not notice a small difference between the checked marking and the template one;
- the manual verification work is monotonous.

Existing projects are based on hardware solutions, for example, on the use of special sensors [4].

Such solutions implement marking recognition, but have a number of important disadvantages:

- Unstable recognition quality, depending on the conditions under which the recognition is performed. Since the production line moves quickly, the necessary conditions for high-quality recognition are often not met;
- Necessity to purchase specialized software to configure sensors.

Thus, such solutions create additional difficulties in operation, which are manifested in the need, in addition to selective manual control of product quality, to control the recognition system itself.

IV. Proposed approach

The proposed approach is to use a pipeline structure of separate neural network modules, each of which solves its own subproblem of marking recognition.

The task of this pipeline is to detect the marking, determine its type and recognize it.

Let us stop on the system architecture in more detail.

A. Formulation of the marking detection problem

As mentioned earlier, in the process of recognizing product markings, additional tasks are solved related to the correct application of such markings. The first task is to determine the presence of marking on the product and the type of marking. The second task is to determine the presence of marking distortions that arise during the printing process, the absence of its parts, etc. And finally, the third task is the actual recognition of the marking, which is carried out in different ways depending on its type (for example, by detecting individual digits that make up the marking, or by recognizing the corresponding Data Matrix code). After completing these tasks, the output information is generated (the date of production of the product, the number in the consignment, etc.) and the correctness of these data is determined in accordance with a predetermined template.

However, it should be noted that in the process of solving these problems, the following problems arise:

- **High speed video stream**. Since the video stream speed is 76 frames per second, the processing time for each frame is about 13 milliseconds. It should be noted that this time is not enough to launch a complex neural network architecture.
- **Impossibility of correct direct detection of digits**. In addition to the digits contained directly in the marking, digits applied to other objects can get into the frame, for example, the production line itself or its parts. In addition, it should be noted that the image enters the neural network with a reduced resolution, which leads to the difficulty of recognizing small objects (digits).
- **Possibility to change the orientation of marking**. During the movement of products along the production line, it is possible to rotate the marking at an arbitrary angle, which leads to a significant degradation in recognition.

The listed problems can be solved architecturally. So, the first of these problems is solved by skipping insignificant frames in which the product is not in the middle of the frame. This allows us to increase the time interval required for image processing by the neural network. Thus, it is necessary to assess the significance of the frame. This can be done with a simple model with a short processing time.

The second problem is solved by decomposing the detection problem into separate subtasks. For example, at the beginning, the product is found, then the marking on the product and, finally, separate digits.

Finally, the third problem is solved by estimating the angle of rotation of the marking in order to present it in the maximum horizontal position.

The presence of several subtasks involves the use of a group of models, each of which does its own part of the work.

B. Recognition module architecture

Based on the previously formulated tasks and problems solved by the recognition module, we list the main subtasks.

1) Assessment of product position
2) Product and marking detection
3) Defining the type of marking
4) Rotate markings for horizontal orientation
5) Marking recognition
6) Assembling the marking and checking it

It should be noted that in our implementation, almost every subtask uses its own neural network model. This allows you to easily modify the system, improve individual modules, change them, and add new ones.

The architecture of the ready-made recognition module is shown in Fig. 3.

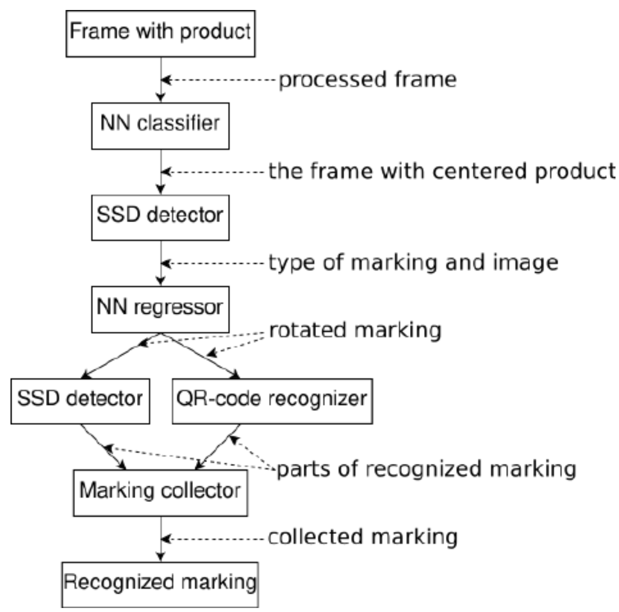

Figure 3. Structure of the marking recognition module

Let us briefly describe the purpose of each individual neural network model and their role in the overall architecture.

The first model is a convolutional classifier that determines the significance of the frame for subsequent analysis. In this case, the most significant is the frame in which the product is closest to the center of the frame. Four main classes of positions were defined as the distance of the product from the center of the frame. Class 1 describes the minimum distance. Only frames of this class are involved in the subsequent stages of processing and analysis. Classes 2 and 3 describe the average and maximum distance. Class 4 is used when there is no product in the frame (empty line).

The architecture of the classifier used is shown in Fig. 4. It consists of 5 layers and has 4 output neurons according to the number of classes that determine the position of the product in the frame. All layers use the ReLU activation function except for the 3rd and last layers. They use linear and softmax activation functions, respectively. Also applies max pooling after the first and second convolutional layers with stride equal to 2.

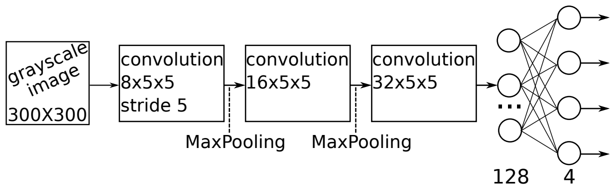

Figure 4. Classifier structure for bottle position estimation [2]

If the frame has been assigned to class 1, it is passed on to the next model.

The second model is a detector and searches for products and markings in the frame. Here, an SSD network [5] based on the MobileNet v1 [6] classifier was chosen as the architecture.

Independent product and marking detection automatically identifies the missing label situation. To do this, it is enough to check the logical condition for the absence of marking in the presence of the product itself in the frame. If a defect is detected, the system notifies the operator about it.

It should be noted that this detector is used to detect different types of markings. Since the SSD model can be used to detect objects of different classes, it was decided to use it to determine the type of markings, which does not require any special analysis procedures.

As a result, if a defect in the marking was not detected, the image of the marking (in its original size) and information about its type are transferred to the next model.

The third model is a regressor used to estimate the angle of rotation of the marking. This model returns the angle by which the marking must be rotated to achieve a horizontal orientation of the image. This conversion improves the quality of subsequent digit recognition. After rotation, the image of the marking is transferred to the model for analyzing the corresponding type of marking (in our implementation, these are models for analyzing a Data Matrix code or alphanumeric marking).

To analyze the alphanumeric marking, the SSD-MobileNet v1 detector is also used, which detects separate digits in the marking. The current implementation uses a non-neural network model to analyze the Data Matrix code. The applicability of a neural network for such an analysis may be the subject of further research.

Further, the assembly of the recognized marking is carried out and the recognition result is returned.

The principle of operation of the neural network component using the example of alphanumeric marking is shown in Fig. 5.

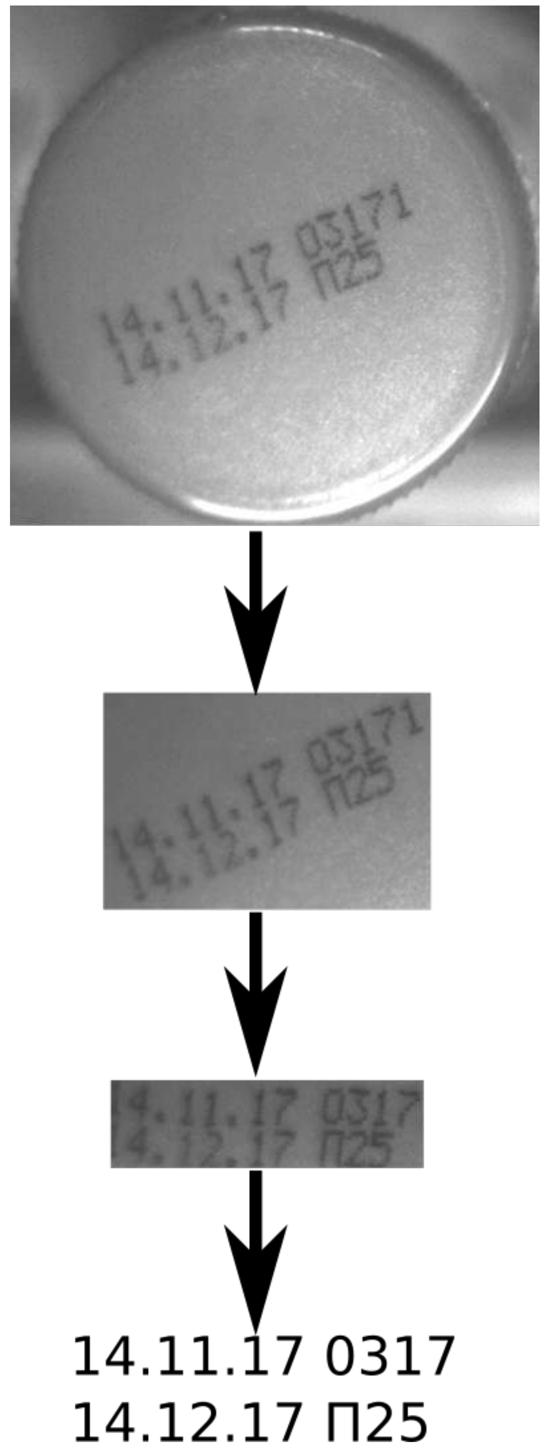

Figure 5. How the system works

C. Training datasets

In the process of preparing neural network models, we used several different training datasets:

- Dataset for training the classifier
- Dataset for training the detector of markings and products (determining the type of marking)
- Dataset for training the rotation angle regressor
- Dataset for training the digit detector

Dataset for training the classifier. To create a dataset, the Faster R-CNN [7] model was used (based on the pretrained ResNet50 [8] classifier). This model has better performance indicators than SSD-MobileNet, but is inferior in speed. It was used to automatically labeling the available data (mainly video files of the production process) according to the distance of the product from the center of the frame. The metric used was the Euclidean distance from the center of the product to the center of the frame. Thus, four classes of images were formed, which were used for the subsequent training of the convolutional classifier. The total sample size was 6189 images, 1303 of which constituted the test dataset.

Dataset for training the detector of markings and products (determining the type of marking). To form this dataset, we used manually labeled images from the classifier training dataset and additional images of generated Data Matrix codes. The total dataset size was 815 images, 163 of which constituted the test dataset.

Dataset for training the rotation angle regressor. When creating this dataset, the images of markings rotated at arbitrary angles were used. The total dataset size was 59385 images, of which 11877 constituted the test dataset. The images of markings obtained from the dataset for training the detector of markings and products were taken as a basis.

Dataset for training the digit detector. To create this selection, a dataset of house numbers SVHN [9] was used, as well as labeled markings. The variant of the SVHN dataset was used, which included 33402 images in the training part and 13068 in the test part. The dataset size of labeled markings was 419 images.

D. Results

After training classifier, the final recognition accuracy was 93.27%.

Both detectors (products/markings and separate digits) were trained on the basis of pre-trained models.

The use of the SSD model allows achieving detection efficiency of 99% (mAP = 0.99) for product detection and marking and 92% (mAP = 0.92) for separate digits. In addition, the processing speed makes it possible to successfully detect markings in the video stream at 76 frames per second. The results of the recognition efficiency of separate digits are presented in the table I.

Table I
EFFICIENCY OF DETECTING INDIVIDUAL CLASSES OF DIGITS

Class label	AP
0	0.9218
1.	0.9107
2.	0.9354
3.	0.9286
4.	0.9265
5.	0.9137
6.	0.9274
7.	0.9167
8.	0.9646
9.	0.8975
mAP	0.92429

The results for the detector of markings and the detector of separate digits are shown in Fig. 6 and 7.

Figure 6. Detected product and marking

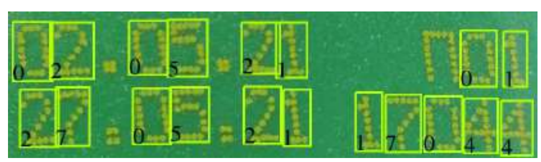

Figure 7. Detected digits in marking

Conclusion

This paper discusses the development of an intelligent component for recognizing product marking based on a neural network approach. The advantage of the proposed solution is the use of the neural networks pipeline, which allows easy switching between individual models. It should be noted the universality of the proposed approach, which is manifested in the simplicity of performing the work on recognizing arbitrary markings on a variety of products. It is enough to integrate a new module with the required functionality and the system will start using it. The models we use are efficient, which makes it possible to operate the system on a fast moving production line.

The direction of further work can be chosen to improve the existing results of recognition of alphanumeric markings, as well as to study the application of a neural network approach for QR code recognition.

ACKNOWLEDGMENT

The author would like to thank the research groups of the Departments of Intelligent Information Technologies of the Belarusian State University of Informatics and Radioelectronics and the Brest State Technical University for their help and valuable comments. In particular, Vladimir Adamovich Golovko, Vladimir Vasilyevich Golenkov, Daniil Vyacheslavovich Shunkevich and Mikhail Vladimirovich Kovalev.

REFERENCES

[1] GS1 DataMatrix – Electronic sign. Available at: https://datamark.by/voprosy-otvety/gs1-datamatrix (accessed 2021, Jun) (In Russian)

[2] V. Golovko, A. Kroshchanka, M. Kovalev, V. Taberko, and D. Ivaniuk “Neuro-Symbolic Artificial Intelligence: Application for Control the Quality of Product Labeling,” Otkrytye semanticheskie tekhnologii proektirovaniya intellektual’nykh system [Open semantic technologies for intelligent systems], 2020, pp. 81-101.

[3] Milk products – Electronic sign [Electronic resource]. Access mode: https://datamark.by/markirovka-molochnoy-produktsii (accessed 2021, Jun) (In Russian)

[4] Omron documentation. Available at: https://robotics.omron.com/browse-documents. (accessed 2021, Jun)

[5] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, A. Berg “SSD: Single Shot MultiBox Detector,” in arXiv, arXiv:1512.02325v5 [cs.CV], Available at: https://arxiv.org/abs/1512.02325 (accessed 2021, Jun).

[6] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” in arXiv, arXiv:1704.04861v1 [cs.CV]. Available at: https://arxiv.org/abs/1704.04861 (accessed 2021, Jun)

[7] S. Ren, K. He, R. Girshick, and J. Sun “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” in arXiv, arXiv:1506.01497v3 [cs.CV]. Available at: https://arxiv.org/abs/1506.01497 (accessed 2021, Jun)

[8] K. He, X. Zhang, S. Ren, and J. Sun “Deep Residual Learning for Image Recognition,” in arXiv, arXiv:1512.03385v1 [cs.CV]. Available at: https://arxiv.org/abs/1512.03385 (accessed 2021, Jun)

[9] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Ng “Reading Digits in Natural Images with Unsupervised Feature Learning,” NIPS Workshop on Deep Learning and Unsupervised Feature Learning, 2011.

Нейросетевой компонент системы распознавания маркировки продукции на производственной линии

Крощенко А. А., Иванюк Д. С.

Данная работа посвящена аспектам разработки нейросетевого компонента системы распознавания маркировки на производственном предприятии ОАО “Савушкин Продукт”. Данная система предназначена для эксплуатации с различными типами маркировок (буквенно-цифровая, Data Matrix) и может быть легко дополнена новыми типами по мере необходимости. Каждый из модулей предложенной системы заменяем и может независимо изменяться и улучшаться. Нейросетевые модели, используемые в текущей реализации системы, отличаются быстротой работы и могут быть с успехом использованы на производственной линии с высокой скоростью конвейера.

Received 09.07.2021