Improving the energy efficiency of sorting centers by identifying objects and digit-letter information with neural networks

Evgeniy Ivliev[*], *Pavel* Obukhov, *Viktor* Ivliev, *Denis* Medvedev, and *Viktor* Martynov

Don State Technical University, Gagarin square 1, 344016 Rostov-on-Don, Russia

Abstract. The article is devoted to the development and analysis of methods of identifying dynamic objects. A neural network with the architecture of SSD MobileNetV2 has been developed to solve the problem of detecting baggage tags and barcodes. Several approaches are considered to solve the problem of identifying digital-letter information: Tesseract, SSD InceptionV2, OpenCV and a convolutional neural network. The efficiency of the methods on real images was checked. It was concluded that electricity consumption can be reduced by 49.43%.

1 Introduction

Despite the widespread automation in some enterprises, it is either absent or is not effective, including not effective in terms of energy consumption.

The relevance of this work is due to the fact that at the moment in most airports, employees of sorting rooms identify baggage tags manually. For comfortable work of a person with manual identification, the room should be equipped with good lighting. It is equally important that it is uninterrupted, meets the quality indicators and lighting standards, and is completely safe. Lighting accounts for a large part of the costs: replacement of burned-out lamps, payment for the services of an employee who is engaged in maintenance, use of lifting equipment, disposal of burned-out lamps.

According to the regulations, the illumination in the sorting and baggage claim areas should be 400 lux. To meet these requirements, such large rooms as the baggage sorting area must be provided with a light flux of ~320000 lumens. For example, if you equip the room with LGT-Prom-Solar-440 lamps with a luminous flux of 55000 lumens and the power consumption of 440 watts, then such lamps will need 6 pieces with a total power consumption of 2640 watts (P_{L400}).

To improve energy efficiency, it is advisable to implement automatic systems with effective subsystems for identifying objects.

* Corresponding author: 123ivliev123@mail.ru

2 Review literature and associated works

In earlier paper [1], an algorithm developed by us to search for barcodes on a baggage tag was used for the localization problem. Later, it was experimentally revealed that this algorithm is not working under various conditions, such as: the distance from the baggage tag to the video camera, illumination, orientation. The algorithm has drawbacks. As a result of the algorithm, an area will be formed that informs about the location of the barcode, but this area does not always correspond to the area of the barcode under various conditions, which negates the further identification of information about the flight of the aircraft.

In the article [2], the authors consider the methods of localization of the baggage tag and identification of the digit-letter information of the IATA airport code. A system for identifying information from a baggage tag based on several neural networks with the SSD InceptionV2 architecture has been developed. These neural networks work with a fairly high accuracy of 82-95% and a speed of 7-10 frames per second. The advantages and disadvantages of using the method of scale-invariant transformation of features for the identification of baggage tags are considered. However, the paper has a significant drawback. To accurately detect the IATA airport code information, the algorithm from [1] is used. The algorithm should rotate the tag image vertically, but there are cases when the tag image is rotated by ±90, ±180 degrees.

3 Materials and methods

As a computing module for implementing the system, Nvidia Jetson TX2 was used. TX2 has a low power consumption of up to 7.5 watts, and the maximum power required at full performance is less than 15 watts ($P_{Jetson\ TX2}$).

To solve the problem of baggage tag localization using neural networks, we considered several different basic neural network architectures available in the TensorFlow Object Detection API, which is an open source platform developed by Google on the basis of TensorFlow and allows you to easily build, train and deploy object detection models.

For the purposes of this study, the following main meta-architectures: Faster R-CNN [3], R-FCN [4], and SSD [5] and feature extractors were considered: Resnet V1 50, Resnet V1 101 [6], Inception V2 [7], Inception Resnet V2 [8] , and Mobilenet V1 [9].

According to the study [10,11], the model SSD MobileNetV1 was chosen because it has a higher identification speed, consumes less memory, and the identification accuracy is comparable to the rest.

Due to time constraints and computational costs, all the experiments presented in this paper use publicly available object detection models that were previously prepared based on the Microsoft COCO dataset [12].

To train the neural network, a training sample consisting of 200 images with baggage tags was created. Data annotation was performed by the LabelImg, which identifies the boundaries of the object of interest and specifies the class to which this object belongs.

For correct recognition of numbers-letter information, the correct display of the tag is required. Image rotation is performed by the functions of the "OpenCV" library:
1. Search for barcode lines by Hough transform [13].
2. Getting the angle of the lines.
3. Affine transformation of the image.

But there are cases when the tag image is rotated by ±90, ±180 degrees. To solve this problem, a convolutional neural network (CNN) was used, which is responsible for classifying the rotation of the baggage tag image. The architecture of this neural network is shown in figure 1.

To solve the problem of recognizing alphanumeric information, 3 methods were proposed.

The first method is the Tesseract program. Tesseract is a software OCR engine that is currently supported by Google, which bought the program in 2006 from HP and opened the source code.

The second method is the neural network SSD MobileNetV2, which has shown good results in localization of baggage tags and barcodes. The neural network was trained on a synthetic training sample consisting of images on which symbols of different fonts were randomly placed. Each symbol was rotated, scaled, distorted, noisy, and discolored for an acceptable variety of characteristics. The proposed method was used to create 8000 images with numbers and letters for 36 different classes.

The third method is the localization of symbols using OpenCV and the classification of symbols by a convolutional neural network. To train the neural network, a database of 76159 monochrome images of numbers and letters of different fonts was created. The neural network architecture is shown in figure 2.
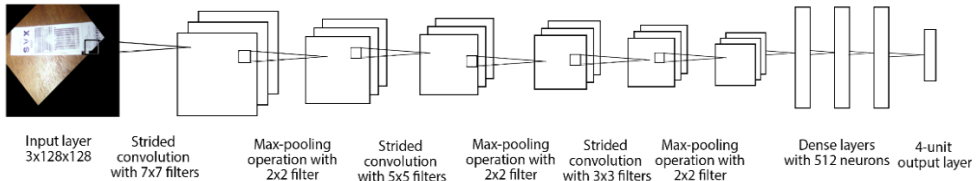

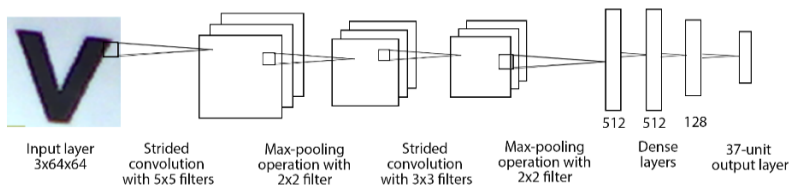**Fig. 1.** Architecture of CNN for classifying the turn of the baggage tag

Fig. 2. Architecture of CNN for Symbol Classification

4 Results and Discussions

As a result of training at 240 eras, the accuracy of the neural network for localizing the baggage tag was 95.3%. Using this neural network, the search area for flight information is narrowed. However, for more precise narrowing of the search area, it is necessary to search the barcode area as the conceived flight information is located above or below the barcode. Therefore, a neural network was created to solve the problem of localizing the barcode. The same image database was used for training, but barcode regions were highlighted during data an-notation. Upon completion of training, the accuracy of localization of the bar-code area was 96.7% when using this neural network. When operating these neural networks, much more computing resources are required than one of them, so to reduce the computational load, it was decided to train one neural network to recognize several classes of objects, rather than teach several neural networks to recognize one class. Upon completion of training, the accuracy of localization of objects was 89% when using this neural network. The result of the neural net-work is shown in figure 3.

The result of the algorithm for turning the baggage tag is shown in figure 4. The accuracy of CNN for classifying the baggage tag rotation was 84%.

To work with Tesseract, we used the Python language, which provides the pytesseract library. However, it was found that the program shows good results only under ideal conditions.

The neural network SSD MobileNetV2 for symbols identification has been trained for 70 epochs. Figure 5 shows the results of this neural network. We can observe that the neural network detects IATA airport codes well, but does not cope well with small symbols.

When using the third method, the localization of characters was performed in the following steps::

1. Capture the area of the image that is above the barcode.
2. Convert the image to grayscale and reverse binarization of the image.
3. Finding contours, figure 6. a

4. To avoid false recognition, filtering is performed by the size of the contours, i.e. the width and height of the contour must be proportionally equal to the size of the letter or number, figure 6. b.

Neural networks for the classification of alphanumeric information were trained for 100 epochs. The accuracy was 93.5%. The result of the work is shown in figure 7.
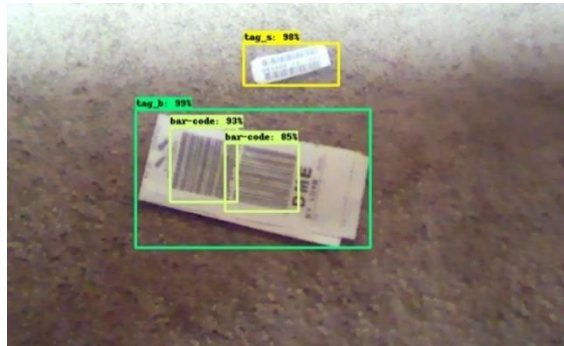

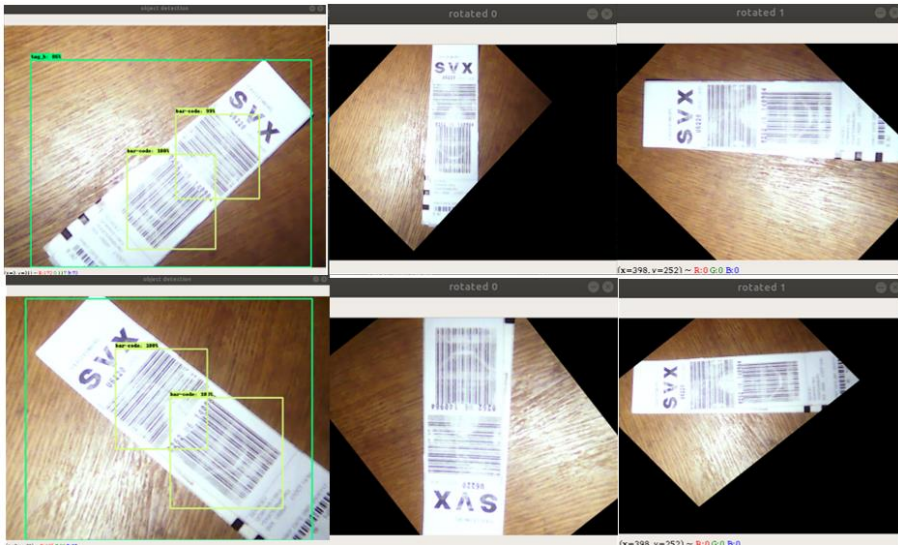**Fig. 3.** Localization of baggage tags and barcodes by neural network

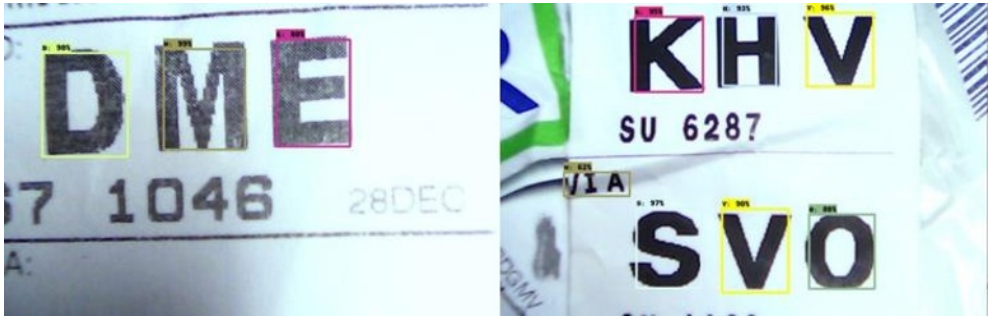**Fig. 4.** Turning the baggage tag

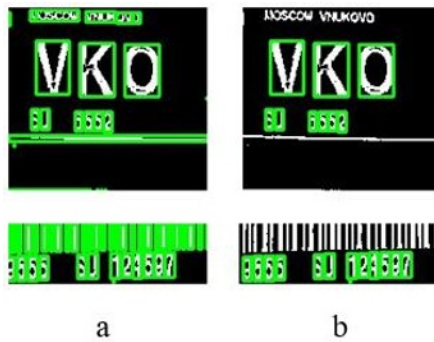**Fig. 5.** Detecting alphanumeric information by SSD MobileNetV2

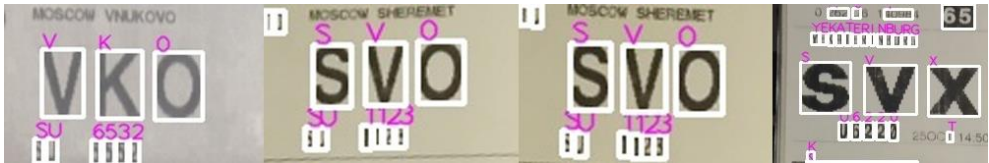**Fig. 6.** a) Search for all contours; b) Sorted contours

Fig. 7. Localization and classification of alphanumeric information

Using this identification system, you can reduce your electricity consumption. Since the identification will be carried out by the system, and not by a person, less light can be used in this area. According to the lighting standards for transport and distribution systems-200 lux, the electricity consumption will be 1320 watts (P_{L200}). Considering the electricity consumption in the first case (without the identification system) and in the second (with the identification system), we can say, we can say that the identification system makes the sorting area more energy efficient by 49.43%(1).

$$k_{ee} = 1 - (P_{L200} + P_{Jetson\ TX2})/P_{L400} \tag{1}$$

5 Conclusion

In the course of this work, a neural network with the SD Mobile Net V2 architecture was developed to solve the problem of detecting baggage tags and barcodes. Several approaches for solving the problem of identifying alphanumeric information are considered. The best result was achieved in the approach where localization was performed by OpenCV tools,

and classification by a convolutional neural network. It was concluded that electricity consumption can be reduced by 49.43%.

References

1. P. S. Obukhov, E. A. Ivliev, V. A. Ivliev, *Identification of alphanumeric information from baggage tag based on neural network a baggage*, Proceedings of the Dynamics of Technical Systems, DTS, 12-13 September 2019, Rostov-on-Don, Russia (2019)

2. E. A. Ivliev, P. S. Obukhov IOP Conf. Ser.: Mater. Sci. Eng. **1029**, 012085 (2021)

3. S. Ren, K. He, R. Girshick, J. Sun, Neural Information Processing Systems, **39**, 1137-1149 (2015).

4. 4 J. Dai, Y. Li, K. He, J. Sun , R-fcn: Object detection via region-based fully convolutional networks. Neural Information Processing Systems, **1**, 379–387, (2016).

5. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, A.C. Berg, SSD: Single shot multibox detector. In European Conference on Computer Vision, Proceedings of the European Conference on Computer Vision (ECCV) 1, 21–37, Springer, (2016).

6. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, **1**, 770–778, (2016).

7. S. Ioffe, C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Proceedings of Machine Learning Research, **37**, 448–456, (2015).

8. C. Szegedy, S. Ioffe, V. Vanhoucke, A. A. Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, AAAI Conference on Artificial Intelligence, cite arXiv:1602.07261 (2017)

9. A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Efficient convolutional neural networks for mobile vision applications, cite arXiv:1704.04861, (2017)

10. J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, Speed/accuracy trade-offs for modern convolutional object detectors, , Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, cite arXiv:1611.10012v3 (2017)

11. E. A. Ivliev, P. S. Obukhov, Assessment of deep neural networks for localization system of baggage tags at airport, European journal of natural history, **1**, (2021).

12. Lin, T.Y. Microsoft COCO: Common objects in context, In ECCV, (2014). DOI:10.1007/978-3-319-10602-1_48

13. D. H. Ballard, Generalizing the Hough transform to detect arbitrary shapes, Pattern recognition, **13(2),** 111-122, (1981)