Neural Network Approach for Estimating the Level and Volume of Liquid in Transparent Containers

Vladimir Golovko
Intelligent Information Technologies Department
Brest State Technical University
Brest, Belarus
gva@bstu.by

Egor Mikhno
Intelligent Information Technologies Department
Brest State Technical University
Brest, Belarus
dzinushi.kun@gmail.com

Aleksandryna Mamyha
Intelligent Information Technologies Department
Brest State Technical University
Brest, Belarus
mamygas@yahoo.comcom

Abstract. **The main purpose of this paper is to represent and investigate a neural network approach for determining the volume of the container and the volume of an opaque liquid ina transparent bounded container.**

To achieve the purpose, we apply two models of neural networks, namely AlexNet and eXnet. We have prepared and created a training dataset.

The results of experiments on determining the level and volume of liquid in transparent bounded containers are presented.

Keywords: training dataset, deep neural network, container volume, liquid level, liquid volume

I. Introduction and the Problem Definition

Advances in computer vision are leading to new practical and research solutions [3, 4]. Currently, an urgent problem inthe field of computer vision is the determination of the volume of opaque liquid in transparent bounded containers. To solve this problem, it is proposed to use a neural network approach [4].

In this paper, we investigate various models for determining the volume of opaque liquid in transparent bounded containers without trademarks. The results obtained allow speaking about the degree of applicability of the used models to the proposed recognition problem.

Most of the research in the field of computer vision is focused on the problems of search (detection [7, 8]) and recognition (classification [9, 10]) of solid physical objects. However, little attention is paid to working with liquid objects[5]. Let us consider the problems that were solved in their article:

1. Calculation of the total volume of the container. It is necessary to calculate the volume of a certain bounded container (50 ml, 200 ml, etc.).

2. Percentage-based estimation of the fullness of the container. It consists in getting information about how full the container is (empty, 10%, 50%, etc.).

3. Comparative estimation of the volume. It is necessary to answer the question of whether it is possible to pour the contents of one container into another one. The planned answer options are: yes, no or I cannot say (because there are opaque containers in the dataset).

4. Estimation based on the placement of the container. It consists in determining the amount of liquid after changing the placement of the container (for example, if it is tilted).

To solve these problems, we used a Container with liquid contents (COQE) dataset with a large number of images, which is reasonable for solving all 4 problems. COQE contains more than 10,000 images of various categories of containers: bottles, glasses, jugs, teapots, etc.

The dataset was formed on an open platform, on which everyone could upload and mark up the necessary images. Random people took photographs of various containers with and without liquid using cameras or mobile phones; the volume of containers was measured using a measuring cup or any other convenient way.

In addition to the usual images, 34 CAD models were also loaded from a three-dimensional warehouse, indicating the degree of correspondence between the CAD models and the containers from the images.

The first three problems are mainly related to estimating the geometry of the container and its contents. The last problem considers the assessment of

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

the behavior of the liquid inside the container, for example, at different tilting angles.

By the time the platform was closed, more than 10,000 marked-up images had been collected, 6386 of which were used for training, 3000 – for testing and 1000 – for checking control.

As the tested models, a modification of the well-known model of the CNN family – ResNet-18 – was considered. We used a variation of the configuration of the model, that was pre-trained on the ImageNet dataset. The context tensor was combined with the conv4 1 layer, which accepts data of size 28 x 28 x 128 as the input one. As a result, the size of the input layer was increased to 28 x 28 x 209. The size of the group sample was 96 images. The resulting architecture was called Context Containers ResNet (CRC).

The classical cross-entropy [6] was used as the error function. The process of training the network was carried out in two stages: the estimation of the volume and contents was conducted separately. We will talk about the results obtained in the following sections.

II. PROPOSED APPROACH

Our problem is to estimate the percentage of the fullness of the container (similar to point 2 discussed above). Models of the CNN family – AlexNet and eXnet – were also used as tested models.

A. Creating datasets

When solving the problem, two datasets were used: one to determine the volume of the container and one to determine the volume of liquid in containers. Data collection was carried out using professional cameras “Canon EOS m50 kit18-150mm” and “Canon PowerShot S5 IS”.

Fig. 1. Examples from the Container Volume Dataset

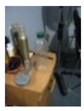
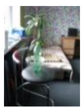
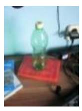
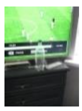
To determine the volume of transparent bounded container, 900 images were collected, 809 of which were allocated for training and 91 – for testing. To determine the volume of liquid in similar conditions, 599 images were collected, 539 of which were allocated for training and 60 – for testing.

Fig. 2. Examples from the Liquid Volume and Level Dataset

We used 3 types of bounded containers: transparent plastic bottles with volumes of 500 ml, 1000 ml and 1500 ml, respectively (excluding brands and subtypes of containers themselves). The investigated volumes of containers as well as the levels and volumes of liquid in containers will be taken into account when marking up the data along with other additional information. The photographs of the bottles were taken with account of different lighting, at different angles, with different sharpness and different backgrounds. The creation of such a context is a necessary criterion for assessing the efficiency of the system under the given conditions.

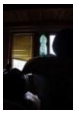
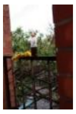
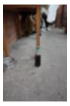
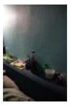
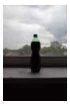
The first dataset consists of transparent plastic bottles with- out liquid and is only needed for the problem of determining the volume of the container itself. For each type of volume, 3 bottles of different producers were chosen. As a result, 9 bottles were taken. All bottles have different shapes except of two: one 1500 ml bottle and one 500 ml bottle have the same shape.

TABLE I. EXAMPLES OF CONDITIONS OF THE CONTAINER VOLUME DATASET

	light	*sharpness*	*corner*	*background*	*room*
500 ml					
1000 ml					
1500 ml					

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

TABLE II. EXAMPLES OF CONDITIONS OF THE LIQUID VOLUME AND LEVEL DATASET

	light	*sharpness*	*corner*	*background*	*room*
0%					
25%					
50%					
75%					
100%					

The second dataset consists of transparent plastic bottles with liquid and is necessary for the problem of determining the volume of liquid in a container as well as for the problem of determining the level of this liquid. 2 bottles of each volume were selected (6 in total), into which an opaque liquid was poured at 4 pre-selected levels: 25%, 50%, 75%, 100% and 0% for empty bottles.

The same containers were used for two datasets.

B. Dataset preprocessing

Before the direct processing the data by the model, a number of transformations of the input data were carried out:

- randomly change the brightness, contrast, intensity and hue of an image;
- resize the input image to the given size;
- rotate the image to an angle;
- crop the given image anywhere;
- horizontally flip the given image randomly with a given probability;
- perform a random perspective transformation of the given image with the given probability.

C. Description of the AlexNet and eXnet models

AlexNet [1] is a universal architecture that can provide high accuracy for complex and large datasets (Fig. 3). AlexNet is one of the most well-known architectures for object detection problems in the field of computer vision. The AlexNet model has about 60 million parameters; the size of the input image – 224x224. In the experiments, we used a pre-trained version of the model on the ImageNet dataset.

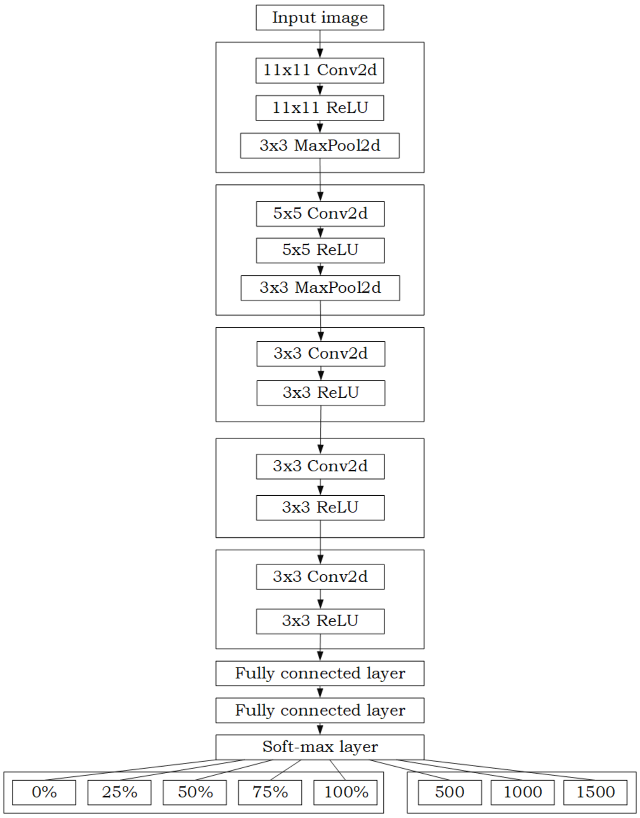


Fig. 3. Architecture of AlexNet

The architecture of the eXnet [2] network (Expression Net) is based on the parallel extraction of objects borrowed from the Inceptions model range but that contains a much smaller number of parameters (eXnet – 4.57 million, AlexNet – 62.3 million, InceptionV4 – over 40 million). Models of this kind provide high performance in systems with limited hardware, and a smaller number of tunable parameters allow high generalization within the framework of the Occam’s razor rule (a simpler explanation of the essence is preferable from two competing theories; this rule states that entities should not multiply unnecessarily). An input image size: 48x48. In the experiments, we used a self-trained version of the model based on standard initialization of weights from Pytorch.

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

An example of architecture is shown in Fig. 4:

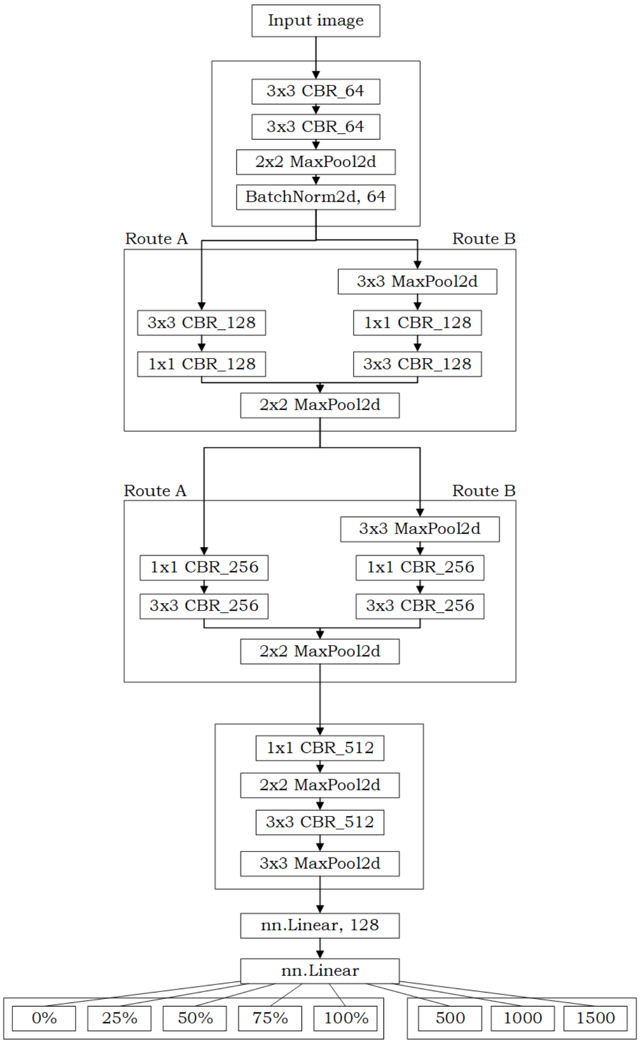


Fig. 4. Architecture of eXnet

III. RESULTS

When comparing the results, validation accuracy was used, at which the values were rounded either to the True Positive side or to the False Negative side.

Six independent experiments were carried out: three for each model at a different number of epochs. The best value for the specified period was taken as the final value. As a result, the AlexNet model showed worse results than the lightweight eXnet model, despite the lack of any pre-training in the case of the last one. In particular, its advantage was observed when assessing the volume of the liquid itself (0.86 versus 0.69).

Based on the test data obtained, we can conclude that the neural network built on the eXnet architecture was well trained, in particular, on a dataset with volumes of containers. The accuracy rate on the second dataset is lower because it contains fewer images and is a more difficult problem for training a neural network. The first dataset has 3 classes,in which different volumes of containers are used. There is no liquid in this dataset. In the second dataset, there are 5 different classes with different volumes of containers (data about containers is used for training in the first dataset) and they also contain an opaque liquid in different volumes.

TABLE III. VOLUME CONTAINER DATASET RESULTS

Epochs	AlexNet	eXnet
30	0.92129	0.93259
50	0.92148	0.92209
100	0.91088	0.94371

TABLE IV. LIQUID VOLUME IN CONTAINER DATASET

Epochs	AlexNet	eXnet
30	0.67794	0.72714
50	0.69378	0.79546
100	0.69432	0.86124

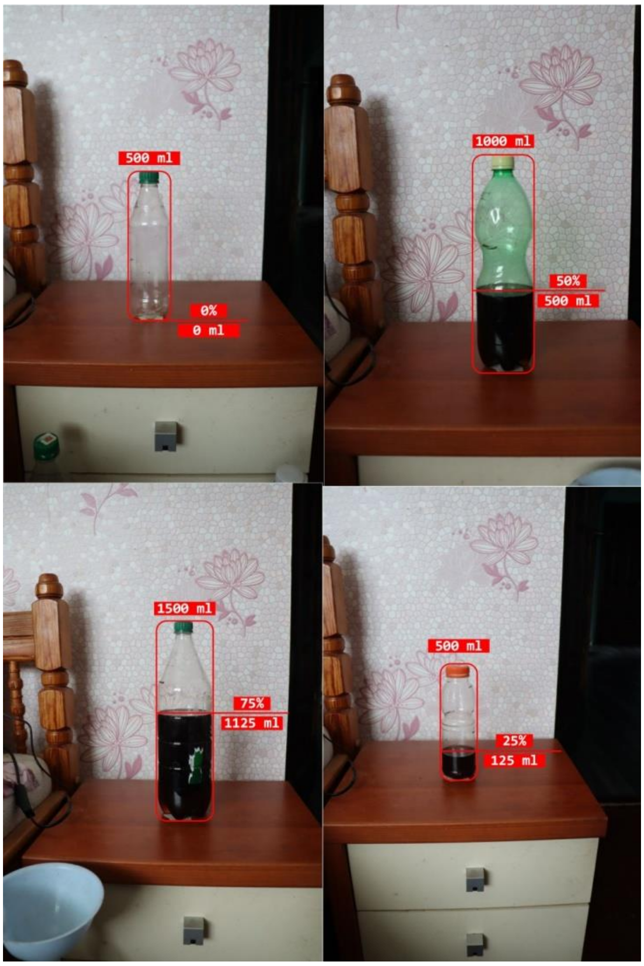


Fig. 5. Result of the system activity

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

The conducted tests allow us to conclude that the data selection and the choice of models were reasonable. To increase theefficiency of the system, it is necessary to increase the amount of data for the conditions described above or to increase the flexibility of the system by increasing the number of conditions themselves.

IV. CONCLUSION

As a result of the conducted studies, a solution was obtained for estimating the volume of the container as well as the level and volume of the liquid in a transparent bounded container with a validation accuracy of 0.94 and 0.86, respectively.

It is generally assumed that the average person copes with the problem by 0.95. The experimental results obtained are still inferior to a human (0.94 and 0.86), so we may continue experiments on this topic in the future.

REFERENCES

[1] Md Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, Mst S. Nasrin, B. C. Van Essen, A. A. S. Awwal, V. K. Asari: "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Ap- proaches." https://arxiv.org/ftp/arxiv/papers/1803/1803.01164.pdf. Ac-cessed 6 Aug 2021

[2] M. N. Riaz: eXnet: "An Efficient Approach for Emotion Recognition in the Wild." https://www.researchgate.net/publication/339324794. Accessed 6 Aug 2021.

[3] V. Golovko, A. Kroshchanka, M. Kovalev, V. Taberko, and D. Ivaniuk "Neuro-Symbolic Artificial Intelligence: Application for Control theQuality of Product Labeling." 2020 International Conference on Open Semantic Technologies for Intelligent Systems, 2020, pp. 81-101.

[4] V. Golovko, A. Kroshchanka, E. Mikhno: "Brands and Caps Labeling Recognition in Images Using Deep Learning." Pattern Recognition and Information Processing, 2019, pp. 35-51.

[5] R. Mottaghi, C. Schenck, D. Fox, A. Farhadi: "See the Glass Half Full: Reasoning about Liquid Containers, their Volume and Content." https://arxiv.org/pdf/1701.02718.pdf. Accessed 6 Aug 2021

[6] J. Brownlee: "A Gentle Introduction to Cross-Entropy for Ma chine Learning." https://machinelearningmastery.com/cross-entropy-for-machine-learning/. Accessed 6 Aug 2021

[7] V. Golovko, A. Kroshchanka, E. Mikhno, M. Komar, A. Sachenko: "Deep convolutional neural network for detection of solar panels". Data- Centric Business and Applications, 2021, pp. 371-389.

[8] V Golovko, A. Kroshchanka, E. Mikhno, S. Bezobrazov, M. Komar, A. Sachenko, I. Shylinska: "Deep convolutional neural network forrecognizing the images of text documents". CEUR Workhop Proceeding,2019, 297-306.

[9] V. Golovko, E. Mikhno, A. Brich, A. Sachenko: "A shallow convo- lutional neural network for accurate handwritten digits classification". International Conference on Pattern Recognition and Information Processing, 2016, pp. 77-85.

[10] A. A. Elngar, M. Arafa, A. Fathy, B. Moustafa, O. Mahmoudm, M. Shaban, N. Fawzy : "Image Classification Based On CNN: A Survey." https://www.researchgate.net/publication/352152485. Accessed 6 Aug2021

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.