The feasibility of applying computer vision methods to automatically determine the composition of heterogeneous disperse systems, using emulsions as a case study, has been considered. This expands the analytical framework, reduces human factor impact on analysis accuracy and reliability, as well as improves processing speed.

During the study, zero-shot segmentation was performed on microscopy images using four different segmenters. The resulting segments were then fitted to circles using a bounding volume (BV) approach. Segmentation effectiveness was evaluated with the Intersection over Union (IoU) metric by comparing results to manually annotated masks provided by an operator.

The average IoU values for the applied segmentation models range from 0.64 to 0.68. Applying the BV technique improves agreement with reference masks; specifically, the average IoU fitted to circles reaches approximately 0.75.

The overall effectiveness of applying the proposed automatic system in the form of a segmentation and bounding volume sequence was determined by analyzing the emulsion droplet diameter distributions. Comparison of the distributions showed that the data obtained using the automatic system are consistent with the operator's data for fractions larger than 15 px. However, the automatic system underestimates the share of fine fractions, which leads to a systematic shift in the integral assessment.

Importantly, it was established that applying the BV method to each individual mask obtained from segmentation is approximately 40–60% faster than analyzing a single combined mask. This analysis of individual masks is also practically more useful in cases involving touching droplets

Keywords: image segmentation, emulsion analysis, computer vision, droplet distribution, droplet diameter

UDC 51-7

DOI: 10.15587/1729-4061.2026.352589

DEVISING OF A SYSTEM FOR ANALYSING THE DISPERSED COMPOSITION OF EMULSIONS USING COMPUTER VISION METHODS

Volodymyr Kosenko
Corresponding author
Assistant*
ORCID: https://orcid.org/0009-0004-1699-5919

Anton Korotynskyi
Doctor of Philosophy (PhD), Senior Lecturer
Department of Automation Hardware and Software**
ORCID: https://orcid.org/0000-0002-6309-5970

Oleksandr Seminskyi
PhD, Associate Professor*
ORCID: https://orcid.org/0000-0002-7081-0899

*Department of Machines and Apparatus of Chemical and Oil Refining Production**
**National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
Beresteiskyi ave., 37, Kyiv, Ukraine, 03056
E-mail: kosenko.volodymyr@lll.kpi.ua

Received 12.12.2025
Received in revised form 21.01.2026
Accepted date 13.02.2026
Published date 27.02.2026

***How to Cite:** Kosenko, V., Korotynskyi, A., Seminskyi, O. (2026). Devising of a system for analysing the dispersed composition of emulsions using computer vision methods. Eastern-European Journal of Enterprise Technologies, 1 (2 (139)), 17–24. https://doi.org/10.15587/1729-4061.2026.352589*

1. Introduction

Emulsions are an important class of dispersed systems, which, due to their special properties, have become widely used. In industry, their application contributes to saving resources, increasing the efficiency of technological processes, and designing innovative products. Consumer goods in emulsion form are distinguished by high quality, structural stability, uniform distribution of active components, controlled release of substances, high indicators of sensory and organoleptic properties.

The most important characteristic of emulsions is the droplet size. Its accurate and rapid determination is necessary for controlling production and product quality, for optimizing technological processes, and for designing products.

Historically, droplet size analysis was based on microscopic observation with manual measurement. The microscopic analysis method has a number of significant advantages, including the possibility of direct visual observation of the emulsion structure, determination of the size and shape of the droplets of the dispersed phase. This makes it possible to obtain qualitative and quantitative information about the degree of dispersion, assessing the stability of the system, the specificity of the droplet structure and their interaction. However, in the conventional version of the measurement using this method, it is laborious, long-term, and dependent on subjective assessments.

Existing methods and means for automatic analysis of emulsions can significantly speed up the process, but require expensive equipment, mainly use indirect measurement methods, and require taking into account their specificity.

Modern achievements in the field of machine vision in combination with classical microscopy open up new opportunities for increasing the speed, accuracy, and reliability of emulsion analysis. Machine vision technologies make it possible to automatically identify droplets in microscopic images and measure them even in the presence of noise and artifacts. Their use does not require the use of additional specific equipment and could be implemented as an extension of the functionality of existing measurement database. At the same time, quite high productivity and reproducibility of measurements are ensured while the human factor is minimized. At the same time, challenges related to image quality, calibration accuracy, and droplet separation require effective solutions.

The above renders special relevance to research aimed at exploring the possibilities of using a combination of microscopy and machine vision methods for emulsion analysis.

2. Literature review and problem statement

Results of the construction of a universal prompt image segmentation model Segment Anything Model (SAM) and the SA-1B dataset are reported in paper [1]; strong zero-shot capabilities on natural images are shown. The applicability to specialized microscopy of small transparent and translucent objects (e.g., emulsion droplets), the stability of detection at low-contrast boundaries, and the calibration of masks for metrological measurements remain unresolved. The limitations are due to the domain shift: the training data almost does not cover emulsions, and the approach itself does not take into account metrology for engineering measurements, focusing on visual scene processing.

In paper [2], a version of SAM is proposed for medical image segmentation for diagnostics, treatment planning, and disease monitoring. It is shown that additional training of the basic model with a specialized dataset significantly improves some segmentation of objects in the application area. The work uses a set of medical images containing 1570263 pairs of images and masks. The effectiveness of using SAM models and their training in the domain of non-medical and engineering applications remains unresolved.

In paper [3], the results of research on the use of deep learning methods for microscopic analysis of images of a wide range of microorganisms are reported. It is shown that deep learning methods successfully solve the tasks of microscopy performed by humans. The exact determination of geometric parameters of objects, such as the size and shape of microorganisms, remains unresolved.

Results of using deep learning for automatic recognition of nanoparticles in microscopy images are described in paper [4]. High accuracy of recognition and determination of particle sizes is shown. The proposed approach helps accelerate research in the field of catalysis and increase the objectivity and accuracy of analysis. The estimation of size distribution with significant particle overlap and quantitative analysis of the shape of objects remain unresolved.

In paper [5], a solution is presented that combines optical microscopy and the Trainable Weka Segmentation tool for determining the droplet size distribution of emulsions was demonstrated and compared with laser diffractometry. It was shown that this approach provides consistent results and allows for the correlation between the size distribution and rheological characteristics (viscosity, turbidity). At the same time, the procedure is significantly dependent on the settings and experience of the operator.

In paper [6], the application of a machine vision is presented, specifically the Circle Hough Transform (CHT) method, for the quantitative analysis of emulsions was demonstrated. The authors developed software to implement CHT with optical images. It was demonstrated that CHT has higher accuracy, reproducibility, and correctness of emulsion droplet identification and size determination compared to manual identification. The analysis accuracy declines when detecting non-circular, deformed droplets and strong contour overlap and high density of placement, as well as the need for careful parameter tuning, remain unresolved.

A deep learning-based microscopic image processing system was designed in paper [7] for automated segmentation and quantitative analysis of microdroplets in colloidal systems. This workflow segments droplets of colloidal systems in microscopic images across a wide range of images and was specifically designed to quantify droplets with a wide range of sizes in a single image. In addition to segmentation, the workflow also integrates the acquisition of quantitative data on all droplets in the image. It should be noted that the cited paper shows how the proposed image analysis system segments and quantifies droplets inside other droplets, as in the case of double emulsions.

In paper [8], the following methods for studying emulsions are considered: visual observation, microscopy, particle size analysis, surface charge analysis, and rheology. Particular attention is paid to microscopic observation. The limitations of microscopy associated with the preparation of the analyzed sample, time consumption, and subjectivity due to the presence of the human factor are also emphasized. However, the authors do not consider modern machine vision methods for quantitative microscopy analysis, in particular automatic droplet segmentation and construction of statistical metrics of droplet distribution. This, in turn, forms a niche for further research: integration of segmentation methods and automated image processing into the emulsion evaluation complex.

In paper [9], the effect of droplet size on the rheological properties of water-in-oil emulsions was investigated. The microstructure was characterized by cryo-electron and optical microscopy; rheology – by amplitude scanning and flow curve measurements. It was found that smaller droplets result in higher yield stresses and elastic modulus, which confirms the need for accurate size determination. At the same time, droplet sizes were measured mainly manually without the use of modern machine vision methods for automated segmentation and error estimation. The lack of such an approach limits the applicability for mass analysis and creates a niche for further research and integration of deep learning methods for automated segmentation.

In paper [10], the rheological behavior of an emulsion is considered to be associated with microstructural changes caused by external loading, which relate the results of rheological characterization to microscopic destabilization phenomena such as coalescence and flocculation, using measurements of droplet size distribution and sequential flow curves. The proposed approach makes it possible to draw conclusions about the progress of microscopic processes in water-in-oil emulsions from rheological data. At the same time, the work does not solve the problem of direct, automated quantitative assessment of microstructure based on images: droplet size distributions are used as an auxiliary tool, and not as a result of a fully formalized image analysis. This limits the capabilities of the method for reproducible assessment of the dispersed composition of emulsions in the general case.

Our review of the literature [1–10] shows significant progress in machine vision methods for microscopic image analysis. Universal models (SAM) have high generalization ability but are limitedly adaptable to specialized emulsion microscopy due to domain shift. The microscopic method of optical microscopy provides acceptable accuracy on idealized images but requires manual intervention. Specialized models demonstrate promising results but their generalization to other modalities remains insufficiently studied.

The task to devise a single automatic system for analyzing heterogeneous systems remains unresolved; therefore, there is a need to develop approaches to the application of machine vision algorithms for automatic determination of the dispersed composition of emulsions.

The above allows us to argue that it is advisable to conduct a study aimed at devising and investigating the possibilities of a system that would combine the universality of SAM models with post-processing methods. This will make it possible to automate the determination of the dispersed composition of emulsions, while ensuring a balance between processing speed and accuracy of results.

3. The aim and objectives of the study

The purpose of our study is to devise a system based on machine vision algorithms for automatic analysis of the composition of heterogeneous disperse systems using the example of emulsions. This will make it possible to expand the analytical and measurement base, reduce the influence of the human factor on the accuracy and reliability of analysis, as well as increase its speed.

To achieve the goal, the following tasks were set:

– to determine the possibility of using SAM models for image analysis and extraction of the necessary visual information; to assess the effectiveness of the models;

– to investigate the effectiveness of using a machine vision method for calculating the disperse composition of emulsions from the obtained segments;

– to conduct a comparative analysis of parameters for the disperse composition of emulsions determined using machine vision and manual measurement.

4. Materials and methods

4. 1. The object and hypothesis of the study

The object of our study is the automated analysis of water-in-oil emulsions microscopic images.

The principal hypothesis assumes that combining a SAM neural network model with the geometric approximation bounding volume (BV) method makes it possible to significantly simplify and accelerate the quantitative analysis of emulsions while ensuring accuracy comparable to the results from manual analysis.

4. 2. Equipment and preparation of heterogeneous disperse systems

The emulsion was prepared in a flow-through rotor-stator apparatus (Ukraine). For this purpose, the apparatus was filled with distilled water, in which 1% (wt.) of the surfactant SLES-70 (Sigma-Aldrich, DE) was dissolved, after which 1% (wt.) of the silicone PDMS-200 (Wacker AK 200, DE) was introduced. The emulsification process was carried out under a circulation mode for 5 min. When this time was reached, samples were taken and used to prepare specimens for microscopic analysis in a standard way.

4. 3. Microscopic imaging

Optical equipment. The dispersed composition of the emulsion was studied using an MBI-1 microscope (56×–1350×) equipped with a Sigeta 2.0 MP digital camera (1/3″ CMOS sensor, Japan). Photographs were acquired in TIFF format at a magnification of 70×. An example image is shown in Fig. 1.

Calibration. The scale in microns was formed using a Sigeta X micrometer objective (1 mm/100 divisions; 0.01 mm/division).

Sample preparation. Sample preparation for microscopic analysis was carried out by regular methods.

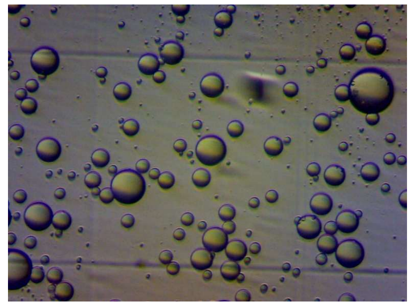


Fig. 1. Photo image of the emulsion sample

As part of our study, 22 images were processed and used for further segmentation and calculation of droplet distribution.

4. 4. Image analysis by the operator

Images were analyzed in the ImageJ 1.54 environment (NIH, USA). The scale "pixel→micron" was set according to the object-micrometer scale, separately for each magnification.

Droplet diameters were determined using the "Straight Line" tool. Only objects with clearly defined boundaries and the absence of optical artifacts that completely fell into the image area were analyzed.

Measurement of the size of emulsion droplets was carried out on 22 images, which provided a total number of droplets in the sample equal to 1525.

To compare the measurement methods, the following were used:

– the average diameter of emulsion droplets according to Sauter

$$d_{3,2} = \frac{\sum_{i=1}^{N} d_i^3}{\sum_{i=1}^{N} d_i^2}, \tag{1}$$

where N is the number of drops; d_i is the diameter of the i-th drop;

– is the minimum diameter of the drop in the sample

$$d_{\min} = \min\left(d_i\right), \quad i = 1 \ldots N, \tag{2}$$

– that is, the smallest drop that has been identified and measured according to accepted criteria.

– the maximum diameter of a drop in the sample

$$d_{\max} = \max\left(d_i\right), \; i = 1 \ldots N, \tag{3}$$

that is, the largest droplet that was identified and measured according to the accepted criteria.

The $d_{\min}$ and $d_{\max}$ values were used as the boundary parameters for diameter distribution and as an additional characteristic of the width of the droplet size range in the sample.

4. 5. Automatic image analysis

Segmentation. To determine image segments corresponding to emulsion droplets, a method of constructing masks based on SAM is considered – an advanced model that provides operational segmentation and is versatile in image analysis tasks.

SAM is able to segment a large number of objects at once in one image, including small objects and even those that were not used in model training before.

In the work of the SAM model, different configurations are used not as objects of comparative analysis but as representative implementations of the segmentation approach. This makes it possible to verify the proposed concept of post-segmentation analysis (SAM + bounding volume) regardless of the specific implementation of the segmenter.

Although SAMs provide correct selection of objects in microscopic images, the results of segmentation in the form of masks are not directly suitable for quantitative analysis of droplet sizes. The unevenness of the contours, local artifacts lead to significant variability of geometric characteristics obtained directly from the masks. Therefore, the work uses the geometric approximation bounding volume – the formation of a closed region that completely describes a geometric object and is often used to increase the efficiency of geometric operations.

Evaluation criterion. The study used pre-trained segmentation models Zigeng_SlimSAM-uniform (ZSU); facebook_sam-vit-large (FSVL); facebook_sam-vit-base (FSVB); facebook_sam-vit-huge (FSVH) under a zero-shot mode, that is, without additional training or domain adaptation on emulsion images. In this regard, the effectiveness of the models within the task was determined by comparing automatically obtained masks with reference masks formed by the operator on the same photographs, with subsequent quantitative assessment of the consistency of the results.

Intersection over Union (IoU) was chosen as the criterion for assessing the similarity of two masks, which is used to describe the degree of overlap of two boundaries. The larger the overlap area, the greater the IoU value.

Automatic image analysis was carried out according to the following algorithm:

1. Reading the microscopic image and reducing it to a standard format supported by the pipeline.
2. Construction of a set of droplet segments using the SAM model in zero-shot mode.
3. Isolation of individual drops and filtering of masks by area and integrity.
4. For each mask of an individual droplet, geometric approximation was performed using the bounding volume method.
5. Based on the bounding volume parameters, equivalent droplet diameters were calculated.
6. A statistical distribution of diameters was formed and generalized characteristics of the dispersed composition were calculated.

4. 6. Computational environment and justification of the choice of methods

Our study was implemented in the Python programming language using the PyTorch, NumPy, Transformers, and OpenCV libraries. Calculations were performed on a Xiaomi Redmi Notebook 14 computer with an Intel Core i5-1135G7 processor (2.4–4.2 GHz) and 16 GB of general-purpose RAM, without the use of an NVIDIA GeForce MX450 graphics processor.

The use of transformer segmentation models is due to their ability to perform zero-shot segmentation without prior training on specialized data sets, which is critically important for the analysis of microscopic images of emulsions with a limited amount of labeled data. Implementation on a CPU makes it possible to assess the practical applicability of the method in a laboratory or research environment without specialized hardware.

5. Results of investigating the application of machine vision algorithms for automatic determination of emulsion composition

The study involves the development of methods that combine fundamental machine learning models with domain-specific retraining and metrological calibration for automated emulsion analysis.

5. 1. Exploring the possibility of using the Segment Anything Model for segmentation of emulsion droplets

Segments of emulsion droplet images were obtained (for example, Fig. 2, *a*), which were compared using IoU (Fig. 2, *c*) with masks generated by the operator (for example, Fig. 2, *b*).

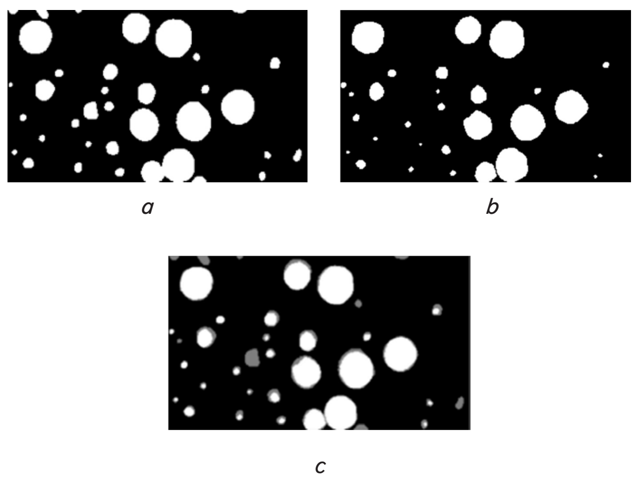


Fig. 2. Samples of segments of an emulsion droplet image: *a* — generated using the Segment Anything Model; *b* — selected by the operator; *c* — visualization of the overlap region of the Segment Anything Model segments and the reference markup of the operator and the corresponding Union, which are used to calculate the Intersection over Union metric

Our results demonstrate that the obtained masks in general have high visual similarity. Larger droplets are quite accurately detected in the images automatically. The results of the study show that the model detects incomplete droplets and their parts at the image boundaries. However, to avoid errors in the analysis, only objects located completely within the image were taken into account.

According to the IoU metric, the accuracy of SAM was assessed on different images; to that end, four segmentation models were used: (ZSU); (FSVL); (FSVB); (FSVH). The results are given in Table 1.

The average IoU value across all test images and models is 0.652, 0.66, 0.675, 0.64, respectively, indicating that the automated image processing has the potential to reliably replace manual analysis.

Table 1

Results of using different SAMs in segmenting emulsion droplets

No.	FSVB	FSVL	FSVH	ZSU
1	0.665187	0.660449	0.674472	0.632947
2	0.618095	0.766832	0.783586	0.644177
3	0.733267	0.730943	0.741375	0.731639
4	0.854219	0.8419	0.8235	0.834092
5	0.740679	0.73552	0.73221	0.709349
6	0.638548	0.680168	0.665226	0.571828
7	0.779153	0.808543	0.79761	0.66245
8	0.693036	0.70625	0.70904	0.648623
9	0.763733	0.667519	0.66372	0.653389
10	0.73829	0.808099	0.813955	0.766751
11	0.52	0.563196	0.571593	0.529406
12	0.581903	0.57889	0.57002	0.564975
13	0.698669	0.712121	0.714593	0.681525
14	0.676901	0.689178	0.691223	0.689334
15	0.598036	0.601821	0.604134	0.598464
16	0.624207	0.646595	0.670168	0.63097
17	0.55	0.566907	0.582876	0.554608
18	0.682175	0.64136	0.660874	0.678001
19	0.628941	0.627336	0.640876	0.613245
20	0.53667	0.542012	0.573234	0.52709
21	0.492845	0.547826	0.569628	0.484345
22	0.529617	0.528677	0.605665	0.525907

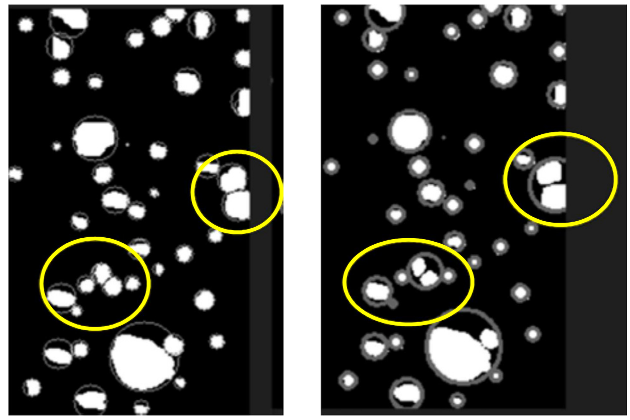
5. 2. Effectiveness of using the machine vision method to calculate the dispersed composition of emulsions from the obtained segments

To effectively determine the dispersed composition of emulsions, it is necessary not only to identify the droplets but also reliably characterize their geometric composition.

To solve this complex task in the automatic analysis of the dispersion of the emulsion, it is proposed to use another machine vision method, namely BV to describe the object by a boundary (in this case, a drop segment – a circle). Then it is possible to calculate the characteristic size of this circle and cluster it within the corresponding geometric segment.

A feature of our implementation of the proposed approach is that the use of BV for each segmented drop separately, in contrast to the processing of one combined mask, reduces the computational complexity and eliminates the effect of sticking of touching drops (Fig. 3). In practical implementation, this leads to a reduction in the total processing time of a single image by 40–60% by reducing the number of geometric operations on complex combined contours: the execution time of a full image analysis cycle is reduced from approximately 5 min to 2–3 min under unchanged computing environment conditions.

To determine the effectiveness of applying box-bounding technique on SAM masks, a study was conducted by comparing the IoU value for the mask obtained as a result of SAM (Fig. 4, *a*) and after bounding volume with circles (Fig. 4, *b*).



a *b*

Fig. 3. Comparison of determining the boundaries of tangential emulsion drops when describing segments with circles: *a* – each mask separately; *b* – generalized mask

The results of using BV to automate emulsion microscopy processing are given in Table 2.

Table 2

Results of using BV for emulsion analysis from images

No.	FSVB	FSVL	FSVH	ZSU
1	0.761678	0.736491	0.746754	0.757941
2	0.778694	0.775424	0.773036	0.762372
3	0.823511	0.740971	0.736529	0.800247
4	0.907258	0.8934	0.8934	0.888822
5	0.773149	0.77227	0.77667	0.753918
6	0.660373	0.657806	0.683089	0.65296
7	0.836742	0.776663	0.762229	0.813661
8	0.707031	0.71002	0.70331	0.700779
9	0.753492	0.692586	0.70197	0.710668
10	0.776877	0.771747	0.76827	0.802108
11	0.6	0.6264	0.606524	0.636422
12	0.780861	0.78428	0.78943	0.792691
13	0.732174	0.746672	0.737586	0.731559
14	0.780912	0.776774	0.772636	0.778498
15	0.740236	0.620906	0.598289	0.709752
16	0.695864	0.693657	0.690435	0.708973
17	0.74	0.780788	0.74953	0.743924
18	0.77281	0.689021	0.781808	0.773643
19	0.812161	0.797553	0.806549	0.797433
20	0.827826	0.830433	0.826483	0.807757
21	0.698297	0.696255	0.698873	0.694523
22	0.77457	0.767005	0.78662	0.752502

The average value of the IoU criterion for all test images and models is 0.75. This indicates that the automated mapping droplet segments with circles has a fairly high accuracy rate relative to the base mask. Therefore, it can be used as a calculation of quantitative indicators based on segmented visual information.

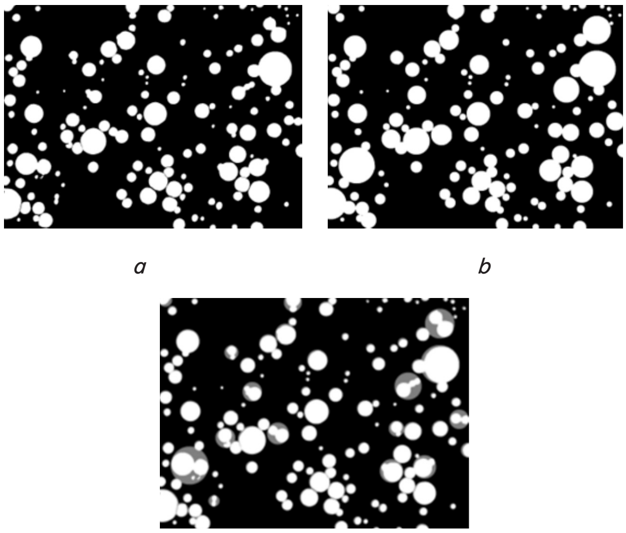


Fig. 4. The droplet masks are obtained as a result of *a* – Segment Anything Model; *b* – Bounding Volume; *c* – visualization of the overlap region of the Segment Anything Model segments and the reference markup of the operator and the corresponding Union, which are used to calculate the Intersection over Union metric.

5. 3. Comparative analysis of parameters for the dispersed composition of emulsions determined by using machine vision and manual measurement

Our study of the effectiveness and feasibility of using the BV combination on SAM masks was carried out by calculating the droplet size distribution obtained using machine vision and comparing them with masks formed by segmenting photographs by the operator. This makes it possible to determine the accuracy of solutions based on machine vision, taking into account accumulated errors, such as the segmenter error in determining the droplet segment and the error in determining the radius of a non-ideally circular segment. The results are represented in the form of normalized emulsion droplet fractions (%) at diameter intervals (the sum of the fractions for each method is 100%). The data for comparison are shown by the plot in Fig. 5.

The operator identifies a large number of small droplets: 62.05% of all measured objects are concentrated in ranges up to 15 px (4.94% in 0–5 px, 34.45% in 5–10 px and 22.66% in 10–15 px). At the same time, SAM+BV models demonstrate significantly worse results in the same ranges – the proportion of droplets is only 19.3–25.6% depending on the model. The most pronounced deviation falls on the interval 5–10 px, which indicates systematic limitations of automatic segmentation in the area of small objects.

In the range of 15–30 px, a systematic increase in particles is observed for SAM models, compared to the manual standard, which indicates a shift in the distribution towards larger diameters. For large fractions above 30 px, the trend of decreasing particles by diameter for SAM and operator data is generally consistent, and the differences between the distributions become less critical. Thus, in medium and large fractions, the automated determination works more stably, and the main area of discrepancy is concentrated in small fractions.

Separately, for each diameter fraction, a model was determined that most closely reproduces the manual standard.

For this purpose, the absolute error by particle in the interval was calculated

$$|\Delta| = |p_{model} - p_{manual}|, \quad (4)$$

where p_{model} and p_{manual} are the shares (%) in the corresponding fraction for the model and manual marking.

According to the results of the calculation for all intervals, the FSVH (7 fractions) and FSVL (6 fractions) models most often demonstrate the minimum error, while ZSU is the best in 2 fractions. The FSVB model did not show the minimum error for any of the considered fractions.

The estimated distribution parameters are given in Table 4. According to the operator's data, d_{32} = 60.097 px was obtained, while for the SAM+BV models from 62.778 px to 69.331 px. That is, automatic methods give a slightly overestimated value of d_{32}, which is consistent with Fig. 5: in SAM, the share of the smallest fractions is underestimated and the share in the middle intervals is increased, due to which the integral estimate is shifted towards larger diameters.

The values of $d_{\min}$ also differ: for SAM models, from 2.000 px to 2.829 px, for the operator, 2.151 px. That is, the determination of this size by SAM models can be considered quite accurate.

The case is somewhat different regarding $d_{\max}$: for SAM models, values from 243.041 px to 247 px were obtained, while for the operator – 225.626 px. That is, all models gave overestimated values, and in all cases by approximately 9.5%.

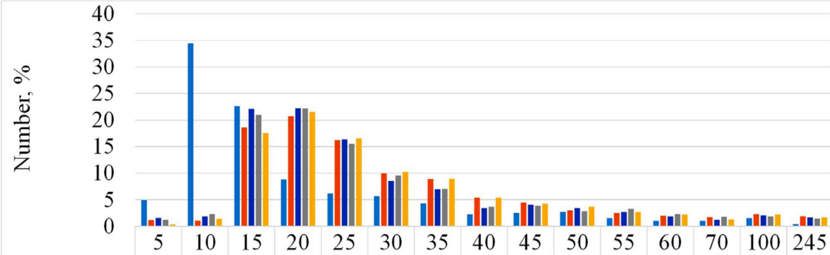
Table 4

Comparison of droplet distribution parameters by diameter (dimensions are given in px)

Criterion \ Estimate	FSVB	FSVL	FSVH	ZSU	Operator
$d_{3,2}$	67.762	67.946	69.331	62.778	60.097
$d_{\min}$	2.829	2.000	2.000	2.829	2.151
$d_{\max}$	243.513	243.178	243.041	247.189	225.626

The droplet diameter distributions are shown in Fig. 5, demonstrating that the operator and SAM results are consistent in ranges above 15 pixels but have significant divergence in small fractions. This is consistent with the findings in Table 4: the system reliably recognizes medium and large droplets but is limited in identifying objects close to its sensitivity threshold.

On average, automatic system allows to achieve ≈ 85% accuracy in determining the size of large and medium droplets, which indicates its suitability for automated analysis of the dispersed composition of emulsions from photographs. The most noticeable deviations are concentrated in small fractions of droplets (1–10 pixels in diameter) and are due to the difficulty of recognizing small circles. However, the recognition of such droplets is important for practical problems of emulsion analysis. Nevertheless, the proposed automatic system is already suitable for emulsion analysis at this stage of development. For this purpose, it is necessary to take into account the specified limitations on the size of droplets and, taking into account the resolution of the photographic equipment used, select the magnification in such a way as to ensure coverage of the entire range of droplet sizes. Another option for using the proposed solution can be the assessment of target droplet fractions.



	5	10	15	20	25	30	35	40	45	50	55	60	70	100	245
Operator	4,94	34,5	22,7	8,83	6,19	5,67	4,28	2,24	2,5	2,7	1,52	1,05	1,05	1,52	0,4
ZSU	1,15	1,06	18,6	20,7	16,2	9,98	8,93	5,37	4,51	2,98	2,5	2,02	1,73	2,3	1,92
FSVH	1,55	1,89	22,1	22,2	16,4	8,52	6,97	3,44	4,04	3,44	2,67	1,89	1,2	2,07	1,64
FSVL	1,2	2,32	21	22,2	15,6	9,54	7,05	3,7	3,87	2,84	3,27	2,32	1,81	1,89	1,46
FSVB	0,37	1,39	17,6	21,5	16,5	10,3	8,96	5,36	4,25	3,7	2,68	2,22	1,29	2,22	1,66

Droplet diameter, pixel

Fig. 5. Normalized droplet diameter distributions and Segment Anything Model+Bounding Volume

Considering the reasonably high accuracy in determining d_{min}, as well as the fact that the relative error in determining d_{32} is approximately 5%, the proposed automatic system can be recommended at this stage of research, especially under conditions when high requirements are not imposed on the accuracy of determining a particular parameter, for example, for comparative assessment of dispersed emulsion compositions.

6. Discussion of results related to the application of machine vision algorithms to automatically determine the composition of heterogeneous dispersed systems

The obtained IoU values (Table 1) of 0.64–0.68 confirm that SAM under a zero-shot mode is able to form droplet masks close to manual marking, but the accuracy is uneven across fractions. The largest deviations are concentrated in the fine-dispersed region, which is in good agreement with Fig. 5: according to the operator, 62.05% of objects are concentrated in fractions up to 15 px, while SAM+BV reproduces only 19.3–25.6% depending on the model. This difference is explained by the fact that in photographs, small droplets have low contrast of the boundary and a small area in pixels. Because of this, they are more often interpreted by the model as background or noise and can also merge/be lost when densely arranged.

The advantage of the applied solution is the combination of two steps: SAM provides fast obtaining of droplet segments, and BV translates the segmentation result into a geometrically interpreted shape (circle), which directly enables the calculation of diameters and construction of distribution. It is practically important that processing each mask separately reduces the risk of "merging" of touching drops into one object (Fig. 3).

In addition, it speeds up the processing by approximately 40–60%. After applying BV, the average IoU values (Table 2) increase to ≈0.75, which indicates the stability of the geometric description of the segments relative to the base masks.

The conventional method of analysis, which involves manual selection of droplet contours and subsequent calculation of the dispersed composition, requires an average of 10–15 min to process one image, depending on its complexity and the number of drops. The proposed automated system makes it possible to reduce the total processing time of one image to 2–3 min without operator participation. In the context of the laboriousness and subjectivity of manual processing, our results demonstrate that the proposed sequence "photograph → SAM-segmentation → BV → droplet sizes" could be used as an automated pipeline for analyzing the dispersed composition.

Unlike the method combining optical microscopy and Trainable Weka Segmentation [5], in which the procedure depends on the classifier settings and operator experience, our result makes it possible to automate the process without the need for manual model training. At the same time, it is the small size fractions that remain the critical zone that most affects the consistency of the distribution and the integral metrics. This is confirmed by the d_{32} indicator (Table 4): for the operator, d_{32} = 60.097 px, while for SAM+BV – 62.778–69.331 px, that is, the automatic system gives a shift towards larger diameters due to underdetection of the small fraction and compensatory growth of particles in the middle intervals.

The limitations of our study are predetermined by the shooting conditions and pixel discretization: the smallest drops have insufficient informativeness of the boundary for reliable segmentation, and the result significantly depends on the quality of the photographs and the selected magnification. An additional limitation is the geometric approximation itself: BV describes the drop as a circle; therefore, for deformed or non-circular segments, an error in the diameter estimation inevitably appears. The disadvantages of the approach in practical application also include the dependence on computational resources and domain shift for fundamental models under a zero-shot mode, which is manifested primarily on small objects.

Further advancement of this work should be directed towards improving the quality of reproduction of small fractions, as well as assessing the uncertainty of diameter measurements for different classes of objects and image acquisition conditions. For example, improving the clarity in determining the boundaries of small droplets by means of shooting and/or specialized post-processing through domain adaptation/retraining of models on emulsion photographs.

7. Conclusions

1. Given the task of applying SAM for segmenting emulsion micrographs, we have found that all considered models give similar results relative to manual marking by IoU: 0.65 (FSVB), 0.66 (FSVL), 0.68 (FSVH), 0.64 (ZSU). The

highest average value was provided by FSVH, which indicates the best agreement of its masks with the operator's standard.

2. Within the task to obtain quantitative characteristics from segments, it is shown that the SAM+BV automatic system makes it possible to move from masks to a geometric description of drops by circles and perform further calculations of the dispersed composition. After fitting the segments to circles, the average agreement with the base masks is ≈ 0.75. Separately, it was found that the analysis of each mask separately is approximately 40–60% faster than the analysis of one combined mask and is practically more useful in cases of touching drops since it reduces the effect of segment merging.

3. As regards the task to compare the dispersed composition to a manual standard, we have determined that the principal discrepancy between the methods is concentrated in the finely dispersed region. According to the operator, the proportion of droplets up to 15 px is 62.05%, while for SAM+BV it is 19.3–25.6%, depending on the model. Starting from fractions above 15px, the nature of the distribution for the models and the standard generally agrees much better, that is, the method is more stable in medium and large fractions. Generalizing the results by integral metrics, it was established that the automated system gives a slightly larger d_{32} compared to the manual standard: for the operator, $d_{32} = 60.097$ px; for the models – 62.778–69.331 px. Such a shift is consistent with the distribution of droplets by diameters and is associated with an underestimation of the proportion of the smallest fractions and a relative shift of the distribution towards larger diameters.

Conflicts of interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study, as well as the results reported in this paper.

Funding

The study was conducted without financial support.

Data availability

Manuscript has associated data in a data repository.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

Authors' contributions

Volodymyr Kosenko: Methodology, Investigation, Resources, Writing – original draft; **Anton Korotynskyi:** Software, Formal analysis, Data Curation, Visualization; **Oleksandr Seminskyi:** Conceptualization, Validation, Writing – review & editing, Supervision, Project administration.

References

1. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L. et al. (2023). Segment Anything. 2023 IEEE/CVF International Conference on Computer Vision (ICCV), 3992–4003. https://doi.org/10.1109/iccv51070.2023.00371
2. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B. (2024). Segment anything in medical images. Nature Communications, 15 (1). https://doi.org/10.1038/s41467-024-44824-z
3. Zhang, Y., Jiang, H., Ye, T., Juhas, M. (2021). Deep Learning for Imaging and Detection of Microorganisms. Trends in Microbiology, 29 (7), 569–572. https://doi.org/10.1016/j.tim.2021.01.006
4. Nartova, A. V., Mashukov, M. Yu., Astakhov, R. R., Kudinov, V. Yu., Matveev, A. V., Okunev, A. G. (2022). Particle Recognition on Transmission Electron Microscopy Images Using Computer Vision and Deep Learning for Catalytic Applications. Catalysts, 12 (2), 135. https://doi.org/10.3390/catal12020135
5. Salum, P., Güven, O., Aydemir, L. Y., Erbay, Z. (2022). Microscopy-Assisted Digital Image Analysis with Trainable Weka Segmentation (TWS) for Emulsion Droplet Size Determination. Coatings, 12 (3), 364. https://doi.org/10.3390/coatings12030364
6. Richards, K. D., Comish, E., Evans, R. C. (2025). Computer vision for high-throughput analysis of pickering emulsions. Soft Matter, 21 (12), 2339–2348. https://doi.org/10.1039/d4sm01252f
7. Saalbrink, J., Loo, T. Y. J., Mertesdorf, J., Xu, P., Pedersen, M. T., Clausen, M. P., Bonilla, J. C. (2025). Quantifying microscopic droplets in colloidal systems through machine learning-based image analysis. Food Hydrocolloids, 166, 111301. https://doi.org/10.1016/j.foodhyd.2025.111301
8. Hu, Y.-T., Ting, Y., Hu, J.-Y., Hsieh, S.-C. (2017). Techniques and methods to study functional characteristics of emulsion systems. Journal of Food and Drug Analysis, 25 (1), 16–26. https://doi.org/10.1016/j.jfda.2016.10.021
9. Tripathi, S., Bhattacharya, A., Singh, R., Tabor, R. F. (2017). Rheological behavior of high internal phase water-in-oil emulsions: Effects of droplet size, phase mass fractions, salt concentration and aging. Chemical Engineering Science, 174, 290–301. https://doi.org/10.1016/j.ces.2017.09.016
10. Marín Castaño, E. P., Leite, R. H. T., de Souza Mendes, P. R. (2021). Microscopic phenomena inferred from the rheological analysis of an emulsion. Physics of Fluids, 33 (7). https://doi.org/10.1063/5.0053408