DOI: 10.63550/ICEIP.2025.70.86.008

Machine Learning Methods Used in the Binarization of Parasitological Micrographs

Natalia Verezubova [1][1], Artur Chekulaev [1][2], Olga Yakovleva [1][3], Natalia Sakovich [2][4]and Irina Verezubova [3][5]

[1] *Moscow State Academy of Veterinary Medicine and Biotechnology named after K.I. Skryabin, Academician Skryabin St., 23, Moscow, 109472, Russia*

[2] *Bryansk State Agrarian University, Sovetskaya St. 2a, Bryansk Region 243365, Russia*

[3] *Moscow City Pedagogical University, 2nd Selskokhozyaistvenny proezd, 4, bldg. 1, Moscow, 129226, Russia*

nverez@mail.ru, tur2205@yandex.ru, yakovleffo@yandex.ru, nasa2610@mail.ru, i.verez@yandex.ru

Keywords : meat raw material analysis, machine learning, binarization.

Abstract: This article discusses the issue of effective use of popular methods of binary image filtering for segmentation of objects in photographs obtained during parasitological studies and veterinary and sanitary examination. A combination of filtering methods and binarization methods not used in classical practice of machine and computer vision is proposed. An optimal structure of a system based on machine learning capable of binary filtering of parasitological microphotographs is proposed. The characteristics of both histograms and the images themselves obtained as a result of parasitological examination are established. In conclusion, statistical processing of the data on histograms and the images themselves obtained during the study is carried out using correlation analysis. Correspondences between the obtained characteristics and the metrics found during the tests are identified in order to determine the influence of image characteristics and segmentation results.

1 INTRODUCTION

The state of the modern world is such that previously non- digitized areas of industry, industry and science, including the agro-industrial complex (AIC), are being transformed under the influence of technological progress.

Digital transformation, which has covered many sectors of the economy, is actively penetrating the agro-industrial complex, creating a new paradigm of “Agriculture 4.0”. This concept involves the integration of advanced technologies such as the Internet of Things (IoT), big data, cloud computing and artificial intelligence into traditional agricultural practices. This leads to the formation of a new technological order, characterized by the automation of processes, increased labor productivity and optimization of resource use.

The use of drones to monitor forest conditions, non-invasive weighing of livestock and other examples of the use of artificial intelligence prove not only its benefits for livestock farming, crop production and ecology, but also the importance of using AI for automation and optimization of the agro-industrial complex.

The introduction of artificial intelligence technologies in the agro-industrial complex demonstrates significant potential in solving complex problems of the industry. Unmanned aerial vehicles equipped with multispectral cameras allow monitoring the condition of crops, identifying foci of plant diseases, determining the need for fertilizers and assessing the condition of forests with an accuracy unachievable with traditional observation methods. Computer vision systems for non-invasive weighing and monitoring of animal health minimize stress in livestock, increase the effectiveness of veterinary

[1] https://orcid.org/0009-0003-4139-8094
[2] https://orcid.org/0009-0005-7424-3106
[3] https://orcid. org/0009-0003-5953-3020
[4] https://orcid.org/0000-0002-6154-4445
[5] https://orcid.org/0009-0002-6265-503X

control and facilitate early detection of diseases. In addition, machine learning algorithms are successfully used to predict crop yields, optimize water consumption for irrigation and plan agricultural work taking into account weather conditions, which significantly increases resource efficiency . agricultural production .

In the modern world, image binarization methods are constantly being improved and updated. Following technical progress, the presented algorithms become more universal, increasing their adaptability. The application of these methods for determining the binarization threshold is very diverse. Their use for both optical character recognition (OCR) and image segmentation has entered the scientific and industrial environments.

The evolution of image binarization methods is characterized by the transition from classical global threshold methods (Otsu , Kapur , Kittler) to adaptive local algorithms that take into account spatial variability of illumination and contrast. Modern approaches include the use of deep convolutional neural networks that can automatically extract relevant features from images and adapt to various shooting conditions. Methods based on wavelet transforms and morphological operations allow for efficient binarization of images with complex backgrounds and uneven illumination. Integrating contextual information and semantic analysis into the binarization process improves the accuracy of segmentation of objects of interest, which is especially important for specialized application areas. Improvement of these methods provides more accurate information extraction from visual data, which is critical for automated decision-making systems in various industries.

The use of such technologies for image segmentation is an extremely relevant topic, in connection with the development of computer vision systems and machine vision, in particular.

Image segmentation is a fundamental task in computer vision, being a necessary stage for subsequent analysis and interpretation of visual data. In the context of the development of machine vision systems, accurate segmentation of objects is of critical importance for such tasks as autonomous navigation, medical diagnostics, industrial quality control and biometric identification. Modern approaches to segmentation, such as semantic segmentation, instance segmentation and panoptic segmentation, allow achieving high accuracy in identifying objects of different classes in images with a complex structure. Deep learning methods, in particular U -Net , Mask R-CNN and DeepLab architectures , demonstrate significant progress in solving segmentation problems in real time, which opens up new opportunities for the practical application of computer vision technologies. The integration of binarization methods with machine learning algorithms creates a synergistic effect, increasing the robustness of segmentation systems to variations in shooting conditions and characteristics of objects of interest.

The use of systems in parasitology can be justified by their applied use. In theory, such models can help reduce the participation of the human factor at the sites of veterinary and sanitary examination of meat raw materials and applied parasitological research.

The introduction of automated image analysis systems in the field of parasitology and veterinary and sanitary examination is a promising direction that can revolutionize traditional research methods. Automatic identification and quantitative assessment of parasitic elements in biological samples can significantly increase the productivity of laboratory research, while ensuring the standardization of procedures and minimizing the subjective factor. Machine vision systems can ensure continuous monitoring of the quality of meat products on conveyor lines, identifying signs of parasitic invasions, such as cysticercosis, trichinosis and echinococcosis, with high sensitivity and specificity. These technologies are especially important in the context of ensuring food safety and preventing zoonotic diseases that pose a threat to public health. Automation of examination processes also helps to reduce labor costs and increase the economic efficiency of veterinary supervision.

Thus, the use of these methods in segmentation of objects in images obtained during parasitological research, remains an insufficiently studied, but promising direction.

Despite the obvious potential, the application of computer vision methods in parasitological research faces a number of specific challenges. The morphological diversity of parasitic objects, their complex texture, variability in size and shape, as well as similarity to surrounding tissues create significant difficulties for automatic segmentation. The limited availability of annotated datasets for training algorithms for recognizing parasitic elements also hinders the widespread implementation of machine learning technologies in this field. However, an interdisciplinary approach combining expertise in parasitology, veterinary science, and computer science can help overcome existing limitations and develop effective methods for automatic analysis of parasitological images. The creation of specialized

algorithms adapted to the specific visualization of parasites and their life stages is a promising area of research with high potential for practical application.

It should be noted that modern methods of image binarization are not suitable for segmentation of objects from the field of parasitology and veterinary and sanitary examination, which may negatively affect the implementation of computer and machine vision technologies in these areas.

Existing image binarization methods developed primarily for standard computer vision tasks demonstrate limited efficiency when working with parasitological specimens and animal tissue samples. Specific characteristics of such images, including low contrast between parasitic elements and surrounding tissues, preparation artifacts, background heterogeneity, and staining variability, require the development of specialized binarization approaches. Traditional methods, such as the Otsu algorithm or adaptive thresholding, often do not provide sufficient accuracy in segmenting parasitic objects, resulting in a high rate of false positives and false negatives. Research shows that combined approaches integrating classical binarization methods with machine learning algorithms and specialized morphological operators can significantly improve the efficiency of parasitological object segmentation. The development of adapted image preprocessing methods that take into account the specifics of parasitological specimens is a prerequisite for the successful implementation of automated diagnostic systems in veterinary practice.

2 SETTING TASKS

Both the tasks and the objectives of the conducted tests are dictated by the relevance of research in the field of computer analysis of parasitological microphotographs. In modern parasitology, automated image processing plays a key role in accelerating diagnostics and increasing the accuracy of parasite identification, which has direct practical significance for clinical medicine and epidemiological monitoring.

The primary goal is to find optimal methods for both binarization for segmentation of objects in parasitological micrographs and methods for preliminary image processing required for this method of binary filtration. Particularly challenging is the heterogeneity of the structure of parasitological preparations, the variability of the morphological characteristics of the objects of study, and the presence of artifacts caused by the process of microscopy and sample preparation. Traditional binarization methods, such as the Otsu method or adaptive threshold algorithms, often demonstrate insufficient efficiency when working with such images, which necessitates the development of specialized approaches.

To implement these tasks, it is necessary to analyze both the characteristics of the images themselves and their histograms. Based on this analysis, it is proposed to develop adaptive algorithms for preliminary filtering, including methods for equalizing histograms, correcting contrast, and suppressing noise of various natures. In addition, it is planned to study the possibility of using multiscale analysis and morphological operators to improve the quality of segmentation of complex-structured objects. The results of this study can find application not only in parasitology, but also in related areas of biomedical visualization that require accurate segmentation of microobjects.

3 ` MATERIALS AND METHODS

The main material of the study will be a dataset containing microphotographs obtained during the veterinary and sanitary examination of meat raw materials affected by trichinellosis. The presented dataset was augmented using rotation augmentation, carried out on the basis of the PyTorch library .

One of the methods used to solve the problem will be the monographic method, used to study previously defined problems and analyze not only individual elements of the problem under study, but also to understand their totality.

The methods used in the study will be binarization methods used both for segmentation of animal cells and for other purposes.

analysis will also involve a binary filtering method, developed to combine the strengths of the segmentation methods under study. The binarization methods will be implemented using the scikit-image library , in Python .

In order to determine the statistical significance of the research results, a method of statistical analysis of the obtained data will be applied, which will be used to analyze the correlation between the parameters of the images and their histograms with the results of the binarization quality.

To obtain objective data on the quality of binarization, an analysis will be applied using metrics such as the Dice coefficient Coefficient), peak signal-to-noise ratio (PSNR), precision , recall .

DOI: 10.63550/ICEIP.2025.70.86.008

Otsu's method . This method of global binary filtering was developed by Japanese scientist Nobuyuki Otsu calculates a threshold that reduces the variance in the two classes (background and object). Thus, we can say that this algorithm is based on clustering.

An obvious disadvantage of the Otsu method is the insufficient stability of systems using this method of binary filtration.

Niblack's method is a local adaptive binarization method based on a sliding window that calculates the threshold using two parameters, the local mean and standard deviation, which are individually calculated for each pixel within its window.

Despite the popularity and ease of use of this method, it is unstable to the appearance of “artifacts” due to the heterogeneity of the background and low contrast of the image.

Adaptive Gaussian Method: This method calculates a threshold for small regions, so that this adaptive threshold is a weighted average of the surrounding pixels, suppressing the high-frequency content of the image.

4 RESULTS AND DISCUSSION

Analysis of the histograms of the images showed a clear bimodal distribution with heavy tails , characteristic of a light object on a dark background. At the same time, negative excess is observed. When measuring the Shannon entropy for these images, an average entropy of more than eleven was found.

Based on the obtained data on the image histograms and their entropy, it can be concluded that parasitological microphotographs of this type are characterized by high noise levels and difficulty in separating the object from the background (low contrast) .

Based on the measurements taken, it can be concluded that such an image requires pre-processing in the form of filtration.

contraharmonic mean filter types were applied .

Also, for more efficient detection of the image threshold taking into account all its features, it was necessary to apply previously unused methods based not only on the use of convolution filters (such as the Sobel operator) with adaptive dilation, but also on the basis of dividing the histogram of the processed image into regions and finding the closest possible threshold using a k-tree.

An integral way to combat histogram homogeneity was to use k -means ++ clustering in such a way that for the first two iterations, the brighter clusters are first filtered out and the centroids in the histograms of subsequent iterations are set as the brightest peaks.

The developed method was compared with classical binarization methods (Otsu method, Niblack method , adaptive Gaussian) in terms of Dice coefficient , recall, recall and peak signal-to-noise ratio (PNSR). The comparison results are presented in the resulting table.

A visualization of the results of different binarization systems is presented in Figure 1.

DOI: 10.63550/ICEIP.2025.70.86.008

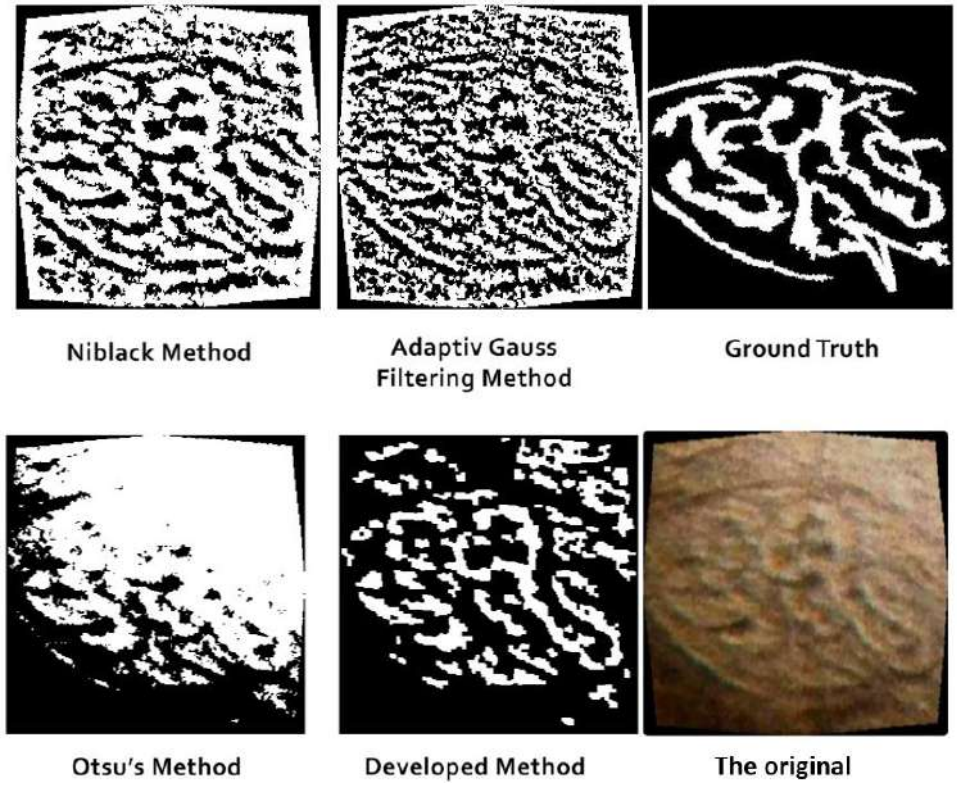

Figure 1: Demonstration of binarization methods.

Table 1 : Results of comparison of binary filtering methods.

Method	Dice coefficient	Precision	Recall	PNSR
Otsu	0.4315	0.2841	0.9227	4,5159
Niblack	0.3522	0.2322	0.8128	3,4459
Adaptive Gaussian	0.3477	0.2429	0,6706	4,1934
Developed method	0.5551	0,6604	0.5053	7,7298

Based on the data in the table, we can conclude that the binarization result is more accurate for the developed method. The Dice coefficient shows that the intersection of ground truth and the binarization result are more accurate than other methods.

The recall index demonstrates high classification accuracy. Using the example of alternative binary filtering methods, it can be understood that these algorithms have a high percentage of noise in the final image and, in general, its incorrect binarization.

The recall metric, which in this case tends to be equal to one for alternative methods, also characterizes other algorithms as less accurate due to the low quality of binarization of image details.

The higher PNSR value of the developed method proves the low noise level of the final data of this method.

Additionally, a correlation analysis was conducted between the Dice coefficient and the Shannon entropy. The analysis was conducted based on the statistical data processing application " Orange " . A multilayer linear perceptron (MLP) with the number of neurons in hidden layers equal to (120,) was chosen as a model for conducting measurements , and SGD was chosen as a solver . The results of the correlation analysis are presented in Table 2.

Table 2: Results of correlation analysis.

Method	Pearson correlation
Niblack	+0 , 823
Otsu	+0,587
Adaptive Gaussian	+0.871
Developed method	-0.135

Based on the data in the table, we can conclude that the low correlation between the Dice coefficient and the Shannon entropy is associated with the low influence of this parameter on the quality of binarization.

Thus, from the results of the test of this simulated system, both its advantages and disadvantages can be deduced.

A distinct advantage of the presented model is its ability to segment objects presented in images included in the created dataset .

Based on the essence of the binarization system, which is a fuzzy logic system using machine learning, we can conclude that during the binary filtering process, the image will either be filtered properly or not filtered at all. This factor is a major drawback of the presented program.

5 CONCLUSION

In this study, an analysis of image histograms was conducted and based on this, an optimal method for binarization of parasitological microphotographs was modeled. This algorithm allows for high-quality segmentation based on binarization of some types of images obtained during veterinary and sanitary examination.

Combining filtering methods with multi-level clustering and active use of undirected graphs allows us to isolate the parasite Trichinella from the general mass of body tissues Spiralis .

Such algorithms can be useful both for pre-processing images to create machine learning datasets , and for independent use to isolate parasitic data, without prior masking, for scientific purposes.

Thus, it can be concluded that the use of machine learning in the analysis of meat raw materials remains an extensive, but still unexplored area.

REFERENCES

Loshkarev , A.S., Timofeev, I.A. , 2021. Research process binarizacii images with use local meaning road . *informatika* . T. 16, No. 6(96). S. 54-65.

Volkov , V. YU., Markelov , OA, Bogachev, MI, 2019. Segmentaciya images i selection object"ektov na basis multi-threshold processing . *News vysshih uchebnyh establishment Russia . Radioelectronics* . Vol. 22, No. 3. P. 24-35.

Afonasenko , AV, 2006. Bystrye morphological transformations for tasks corrections i transformations binary images . *News Tomsk politekhnicheskogo University* . Vol. 309, No. 8. P. 122-125.

Hotilin , MI , Paringer , RA, Rycarev , IA, Kravcova , NS, 2018. Development i analize methods allocations object"ektov images . Information technology i nanotechnology : Sbornik trudov ITNT-2018, Samara, April 24–27 , 2018 . Samarskij national investigator university names academician S. P. Koroleva . – Samara: Predpriyatie «Novaya tekhnika », 2018. S. 2581-2589.

Batko , Yu., Batryn , N., Melnyk , G., 2017. Development of algorithms for biomedical image segmentation based on preliminary markup and texture attributes. *Eastern-European Journal of Enterprise Technologies* Vol. 6, No. 4(90). P. 35-44.

Gagarina , LG, Koldaev , VD, 2014. Methodological approaches to the circuit segmentacii images in automatic manufactured systems . *News vysshih uchebnyh establishments . Electronics* . No. 4(108). P. 64-72.

A threshold selection method from gray-level histograms/N. Otsu. IEEE Trans. Sys., Man., Cyber.1979, 9: 62-66.

State certification registrations programs for EVM No. 2024616826 Russian Federation . Program for binarizacii adaptive images porogom , vychislyaemym po method Nibleka : No. 2024615679: posted . 15.03.2024: published . 25.03.2024 / ES SHtrunova ; posted ' Federal'noe state byudzhetnoe educational education vysshego education « Ryazan state radio engineering university imeni VF Utkina ". – EDN MMGRRM.

Yuldashev , ZM, 2013. Obrabotka medical low-contrast images . *Biomedical radioelektronika* . No. 1. S. 39-41.

Chernyj , SA Locally-adaptive filtering astonished images na basis diskretnogo cosine transformations in the MATLAB environment / SA Chernyj // Molodezhnyj scientific and technical vestnik . – 2012. – No. 5. – S. 22. – EDN PMEJAT.

Volkov , V. YU., 2017. Adaptivnoe selection melkih object"ektov na cifrovyh images . *News vysshih uchebnyh establishments Russia . Radioelectronics* . No. 1. P. 17-28.